Amazon EMR Guide de version

Amazon EMR

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon EMR: Amazon EMR Guide de version

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Les marques commerciales et la présentation commerciale d’Amazon ne peuvent pas être utilisées en relation avec un produit ou un service extérieur à Amazon, d’une manière susceptible d’entraîner une confusion chez les clients, ou d’une manière qui dénigre ou discrédite Amazon. Toutes les autres marques commerciales qui ne sont pas la propriété d’Amazon appartiennent à leurs propriétaires respectifs, qui peuvent ou non être affiliés ou connectés à Amazon, ou sponsorisés par Amazon.

Table of Contents

À propos des versions Amazon EMR 1
Support standard 2
Comprendre les versions d'Amazon EMR 2
Versionnage des versions 3
Politique de prise en charge 4
Considérations 7
Versions Amazon EMR 7.x 9
Versions des applications dans les versions 7.x d'Amazon EMR 10
emr-7,9.0 10
emr-7,8.0 50
emr-7,7.0 96
emr-7.6.0 149
emr-7.5.0 227
emr-7.4.0 303
emr-7.3.0 380
emr-7.2.0 458
emr-7.1.0 577
emr-7.0.0 637
Versions Amazon EMR 6.x 707
Versions des applications dans les versions 6.x d'Amazon EMR 709
emr-6.15.0 709
emr-6.14.0 775
emr-6.13.0 839
emr-6.12.0 911
emr-6.11.1 988
emr-6.11.0 1063
emr-6.10.1 1116
emr-6.10.0 1191
emr-6.9.1 1248
emr-6.9.0 1312
emr-6.8.1 1375
emr-6.8.0 1438
emr-6.7.0 1501
emr-6.6.0 1576

emr-6.5.0 1654
emr-6.4.0 1683
emr-6.3.1 1718
emr-6.3.0 1746
emr-6.2.1 1780
emr-6.2.0 1808
emr-6.1.1 1842
emr-6.1.0 1863
emr-6.0.1 1889
emr-6.0.0 1907
Versions Amazon EMR 5.x 1929
Versions des applications dans les versions 5.x d'Amazon EMR 1933
emr-5.36.2 1933
emr-5.36.1 1977
emr-5.36.0 2038
emr-5.35.0 2088
emr-5.34.0 2116
emr-5.33.1 2142
emr-5.33.0 2172
emr-5.32.1 2195
emr-5.32.0 2221
emr-5.31.1 2250
emr-5.31.0 2269
emr-5.30.2 2293
emr-5.30.1 2313
emr-5.30.0 2335
mr-5.29.0 2359
emr-5.28.1 2378
emr-5.28.0 2397
emr-5.27.1 2417
emr-5.27.0 2435
emr-5.26.0 2455
emr-5.25.0 2475
emr-5.24.1 2495
emr-5.24.0 2512
emr-5.23.1 2531

emr-5.23.0 2548
emr-5.22.0 2567
emr-5.21.2 2587
emr-5.21.1 2604
emr-5.21.0 2622
emr-5.20.1 2641
emr-5.20.0 2658
emr-5.19.1 2678
emr-5.19.0 2695
emr-5.18.1 2713
emr-5.18.0 2730
emr-5.17.2 2747
emr-5.17.1 2764
emr-5.17.0 2780
emr-5.16.1 2798
emr-5.16.0 2814
emr-5.15.1 2831
emr-5.15.0 2847
emr-5.14.2 2864
emr-5.14.1 2880
emr-5.14.0 2896
emr-5.13.1 2914
emr-5.13.0 2930
emr-5.12.3 2946
emr-5.12.2 2961
emr-5.12.1 2977
emr-5.12.0 2992
emr-5.11.4 3009
emr-5.11.3 3025
emr-5.11.2 3040
emr-5.11.1 3056
emr-5.11.0 3071
emr-5.10.1 3087
emr-5.10.0 3103
emr-5.9.1 3120
emr-5.9.0 3135

emr-5.8.3 3152
emr-5.8.2 3167
emr-5.8.1 3182
emr-5.8.0 3197
emr-5.7.1 3213
emr-5.7.0 3228
emr-5.6.1 3244
emr-5.6.0 3259
emr-5.5.4 3274
emr-5.5.3 3289
emr-5.5.2 3304
emr-5.5.1 3319
emr-5.5.0 3334
emr-5.4.1 3350
emr-5.4.0 3365
emr-5.3.2 3380
emr-5.3.1 3395
emr-5.3.0 3410
emr-5.2.3 3425
emr-5.2.2 3440
emr-5.2.1 3454
emr-5.2.0 3470
emr-5.1.1 3485
emr-5.1.0 3500
emr-5.0.3 3515
emr-5.0.2 3529
emr-5.0.1 3543
emr-5.0.0 3557
Versions Amazon EMR 4.x 3572
Versions des applications dans les versions 4.x d'Amazon EMR 3574
Différences de version 3574
emr-4.9.6 3621
emr-4.9.5 3636
emr-4.9.4 3650
emr-4.9.3 3664
emr-4.9.2 3679

emr-4.9.1 3693
emr-4.8.5 3708
emr-4.8.4 3722
emr-4.8.3 3736
emr-4.8.2 3751
emr-4.8.1 3766
emr-4.8.0 3780
emr-4.7.4 3794
emr-4.7.3 3808
emr-4.7.2 3821
emr-4.7.1 3835
emr-4.7.0 3848
emr-4.6.1 3863
emr-4.6.0 3875
emr-4.5.0 3889
emr-4.4.0 3900
emr-4.3.0 3913
emr-4.2.0 3924
emr-4.1.0 3934
emr-4.0.0 3943
Versions AMI 2.x et 3.x d' 3950
Création d’un cluster 3951
Installation des applications 3954
Personnalisation des configurations 3954
Hive 3962
HBase 3973
Pig 3986
Spark 3992
S3 DistCp 3996
Nouveautés 3999
Amazon EMR 7.9.0 3999
Amazon EMR 6.15.0 4010
Amazon EMR 5.36.2 4046
Compatibilité SigV4 4066
Approche visant à atténuer le CVE-2021-44228 4066

Solution d'action d'amorçage Amazon EMR pour Log4j CVE-2021-44228 et CVE-2021-45046 4067
Questions fréquentes (FAQ) 4075
Archivage 4077
Version 6.14.0 4077
Version 6.13.0 4112
Version 6.12.0 4154
Version 6.11.1 4201
Version 6.11.0 4246
Version 6.10.0 4272
Version 6.9.0 4300
Version 6.8.0 4334
Version 6.7.0 4370
Version 6.6.0 4418
Version 5.35.0 4469
Version 5.34.0 4473
Version 6.5.0 4476
Version 6.4.0 4478
Version 5.32.0 4487
Version 6.2.0 4492
Version 5.31.0 4501
Version 6.1.0 4506
Version 6.0.0 4513
Version 5.30.1 4518
Version 5.30.0 4523
Version 5.29.0 4529
Version 5.28.1 4530
Version 5.28.0 4532
Version 5.27.0 4534
Version 5.26.0 4537
Version 5.25.0 4540
Version 5.24.1 4543
Version 5.24.0 4544
Version 4546
Version 5.22.0 4548
Version 5.21.1 4552

Version 5.21.0 4553
Version 5.20.0 4556
Version 5.19.0 4559
Version 5.18.0 4561
Version 5.17.1 4562
Version 5.17.0 4562
Version 5.16.0 4564
Version 5.15.0 4565
Version 5.14.1 4566
Version 5.14.0 4566
Version 5.13.0 4569
Version 5.12.2 4569
Version 5.12.1 4570
Version 5.12.0 4570
Version 5.11.3 4572
Version 5.11.2 4572
Version 5.11.1 4572
Version 5.11.0 4573
Version 5.10.0 4574
Version 5.9.0 4576
Version 5.8.2 4578
Version 5.8.1 4578
Version 5.8.0 4579
Version 5.7.0 4581
Version 5.6.0 4582
Version 5.5.3 4582
Version 5.5.2 4583
Version 5.5.1 4583
Version 5.5.0 4583
Version 5.4.0 4585
Version 5.3.1 4586
Version 5.3.0 4586
Version 5.2.2 4587
Version 5.2.1 4587
Version 5.2.0 4588
Version 5.1.0 4589

Version 5.0.3 4590
Version 5.0.0 4590
Version 4.9.5 4592
Version 4.9.4 4592
Version 4.9.3 4592
Version 4.9.2 4593
Version 4.9.1 4593
Version 4.8.4 4594
Version 4.8.3 4594
Version 4.8.2 4595
Version 4.8.0 4595
Version 4.7.2 4596
Version 4.7.1 4597
Version 4.7.0 4597
Version 4.6.0 4599
Version 4.5.0 4601
Version 4.4.0 4601
Version 4.3.0 4603
Version 4.2.0 4605
Configuration des applications 4607
Configuration des applications lorsque vous créez un cluster 4610
Fournir une configuration dans la console lorsque vous créez un cluster 4610
Fournir une configuration à l'aide de l' AWS CLI option lorsque vous créez un cluster 4610
Fournir une configuration à l'aide du kit SDK Java lors de la création d'un cluster 4611
Reconfigurer un groupe d'instances dans un cluster en cours d'exécution 4612
Considérations à prendre en compte lors de la reconfiguration d'un groupe d'instances 4613
Reconfigurer un groupe d'instances dans la console 4616
Reconfigurer un groupe d'instances à l'aide de la CLI 4617
Reconfigurer un groupe d'instances à l'aide du SDK Java 4622
Dépannage 4623
Stocker les données de configuration sensibles dans AWS Secrets Manager 4626
Créer un secret 4627
Accordez à Amazon EMR l'accès pour récupérer le secret 4627
Utiliser le secret dans une classification de configuration 4627
Mettre à jour la valeur d'un secret 4629
Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique 4629

Considérations 4629
Remplacer la JVM 4632
Ports de service 4635
Utilisateurs de l'application 4637
Vérification des dépendances à l'aide du référentiel d'artefact 4638
Système de fichiers EMR (EMRFS) 4641
Vue cohérente 4643
Activation de la vue cohérente 4647
Comprendre comment la vue cohérente d'EMRFS effectue le suivi des objets dans Amazon S3 4649
Logique des nouvelles tentatives 4650
Métadonnées de la vue cohérente EMRFS 4652
Configuration des notifications de cohérence pour CloudWatch et Amazon SQS 4655
Configuration de la vue cohérente 4657
Référence des commandes CLI EMRFS 4662
Autorisation d'accès aux données EMRFS dans Amazon S3 4674
Création d'un fournisseur d'informations d'identification personnalisées pour les données EMRFS dans Amazon S3 4674
Gestion du point de AWS Security Token Service terminaison par défaut 4676
Spécification du chiffrement Amazon S3 en utilisant les propriétés EMRFS 4677
Utilisation de AWS KMS keys pour le chiffrement EMRFS 4679
Chiffrement côté serveur sur Amazon S3 4680
Chiffrement côté client sur Amazon S3 4682
Chiffrement côté client Amazon S3 avec S3A 4691
Configuration de CSE-KMS 4692
Configuration de CSE-CUSTOM 4692
Propriétés du chiffrement côté client Amazon S3 avec S3A 4694
CloudWatch Agent Amazon 4696
Créer un cluster 4698
Autorisations requises 4698
Point de terminaison requis 4699
Créer le cluster 4699
Métriques par défaut 4700
Configuration 4703
Amazon EMR R 7.1.0 4703
Amazon EMR 7.0.0 4738

Considérations 4744
Historique 4744
Delta Lake 4750
Introduction 4750
Utilisation des clusters de Delta Lake 4751
Delta Lake avec Flink 4751
Delta Lake avec Trino 4756
Delta Lake avec Spark 4757
Delta Lake avec Spark et Glue 4764
Considérations 4764
Historique 4765
Flink 4767
Création d'un cluster avec Flink 4770
Configuration de Flink 4771
Hive et Glue 4771
Fichier de configuration 4773
Plusieurs nœuds primaires 4775
Taille du processus de mémoire 4776
Taille du fichier de sortie du journal 4776
Java 11 4778
Tâches Flink 4782
Démarrez une application Flink YARN en tant qu'étape sur un cluster de longue durée 4783
Soumettre du travail à une application Flink existante sur un cluster de longue durée 4785
Soumission d'une tâche Flink transitoire 4786
Shell Flink Scala 4788
Interface utilisateur Flink 4789
Outil de mise à l’échelle automatique Flink 4790
Présentation 4790
Considérations 4791
Activation de l’outil de mise à l’échelle automatique 4791
Configurations 4793
Optimisation des temps de redémarrage 4799
Récupération locale des tâches 4799
Points de contrôle incrémentiels 4800
Récupération précise 4801
Redémarrage combiné 4802

Flink avec Zeppelin 4803
Introduction 4803
Prérequis 4803
Configurer Zeppelin-Flink sur un cluster EMR 4804
Exécuter des tâches Flink avec Zeppelin-Flink sur un cluster EMR 4805
Historique des publications de Flink 4810
Notes de mise à jour de Flink par version 4851
Ganglia 4854
Création d'un cluster à l'aide de Ganglia 4856
Affichage des métriques Ganglia 4857
Métriques Hadoop et Spark dans Ganglia 4858
Historique des versions de Ganglia 4859
Hadoop 4905
Configuration de Hadoop 4907
Configuration de la tâche 4908
Paramètres de configuration de démon Hadoop 5342
Configuration HDFS 5668
Chiffrement transparent dans HDFS sur Amazon EMR 5669
Configuration du chiffrement transparent HDFS 5670
Considérations pour le chiffrement transparent HDFS 5673
Serveur de gestion des clés Hadoop 5673
Chiffrement transparent HDFS sur les clusters EMR avec plusieurs nœuds principaux 5677
Création ou exécution d'une application Hadoop 5679
Créer des fichiers binaires à l'aide d'Amazon EMR 5679
Traitement des données avec Streaming 5682
Traitement de données à l'aide d'un fichier JAR personnalisé 5688
Lire les objets restaurés 5691
Exemples 5692
Considérations 5693
Activer la détection non uniforme des accès à la mémoire pour les conteneurs YARN 5693
Emballage en poubelle YARN 5695
Activez la fonctionnalité 5696
Considérations 5696
Historique des versions de Hadoop 5696
Notes de mise à jour de Hadoop par version 5762
HBase 5776

Création d'un cluster avec HBase 5780
Création d'un cluster à HBase l'aide de la console 5780
Création d'un cluster à HBase l'aide du AWS CLI 5780
HBase sur Amazon S3 (mode de stockage Amazon S3) 5781
Activation HBase sur Amazon S3 5783
Utilisation d'un cluster réplica en lecture 5784
HFile Suivi permanent 5786
Suivi des fichiers du magasin 5788
Considérations opérationnelles 5789
Journaux d'écriture anticipée (WAL) pour Amazon EMR 5794
Espaces de travail WAL 5795
Autorisations requises 5796
Activation du WAL 5797
Restauration à partir de WAL 5800
Configurations de la sécurité 5801
En utilisant AWS PrivateLink 5802
Tarification et indicateurs du WAL 5803
Marquage des espaces de travail WAL 5805
Réplication entre clusters EMR WAL 5806
Considérations et disponibilité 5808
Référence de la CLI EMRWAL 5810
Utilisation de la HBase coque 5813
Créer une table 5814
Placement d'une valeur 5814
Obtention d'une valeur 5814
Supprimer une table 5814
Accédez aux HBase tables avec Hive 5814
Utilisation de HBase snapshots 5816
Création d'un instantané à l'aide d'une table 5816
Suppression d’un instantané 5817
Affichage des informations d'instantané 5817
Exportation d'un instantané vers Amazon S3 5817
Importer un instantané depuis Amazon S3 5818
Restaurer une table à partir d'instantanés dans le shell HBase 5819
Configuration HBase 5820
Changements d'allocation de mémoire dans YARN 5821

HBase numéros de port ... 5822
HBase paramètres du site pour optimiser ... 5822
Considérations sur les performances ... 5824
Afficher l'interface HBase utilisateur ... 5826
Afficher les fichiers HBase journaux ... 5828
Moniteur HBase avec Ganglia ... 5829
Surveillance de l'EMR HBase avec Amazon CloudWatch ... 5831
Configurer des métriques ... 5833
Utilisation de la destination des métriques ... 5836
Migration depuis les versions précédentes HBase ... 5837
Migration vers Amazon EMR version 7.4.0 ou ultérieure ... 5837
HBase historique des versions ... 5838
HCatalog ... 5917
Création d'un cluster avec HCatalog ... 5918
En utilisant HCatalog ... 5919
Désactivation de l'écriture directe lors de l'utilisation HCatalog HStorer ... 5920
Création d'une table à l'aide de la HCat CLI et utilisation de ces données dans Pig ... 5920
Accès à la table à l'aide de Spark SQL ... 5921
Exemple : création d'une HCatalog table et écriture dans celle-ci avec Pig ... 5923
HCatalog historique des versions ... 5924
Hive ... 5996
Différences et considérations pour Hive sur Amazon EMR ... 5999
Différences entre Apache Hive sur Amazon EMR et Apache Hive ... 5999
Différences Hive entre les versions Amazon EMR 4.x et 5.x ... 6000
Fonctionnalités supplémentaires de Hive sur Amazon EMR ... 6001
Configuration d'un métastore externe pour Hive ... 6006
Utilisation du catalogue de données AWS Glue comme métastore pour Hive ... 6007
Utilisation d'une base de données MySQL externe ou d'Amazon Aurora ... 6014
Utilisation du pilote Hive JDBC ... 6017
Amélioration des performances de Hive ... 6020
Activation du validateur optimisé Hive EMRFS S3 ... 6020
Utilisation de S3 Select ... 6021
Optimisation MSCK ... 6024
Utiliser Hive LLAP ... 6025
Activer LLAP ... 6025
Démarrer LLAP sur votre cluster ... 6027

Vérifier l'état de LLAP ... 6027
Démarrer ou arrêter LLAP ... 6028
Redimensionner le nombre de démons LLAP ... 6028
Chiffrement dans Hive ... 6028
Chiffrement modulaire du parquet dans Hive ... 6028
Chiffrement en transit dans HS2 ... 6032
Historique des versions de Hive ... 6034
Notes de mise à jour de Hive par version ... 6107
Hudi ... 6203
Comment fonctionne Hudi ... 6205
Comprendre les types de stockage des jeux de données : copie en écriture ou fusion en lecture ... 6206
Enregistrement d'un jeu de données Hudi auprès de votre métastore ... 6207
Considérations et restrictions ... 6207
Créez un cluster avec Hudi installé ... 6209
Travailler avec un jeu de données Hudi ... 6210
Initialisation d'une session Spark pour Hudi ... 6214
Écrire dans un jeu de données Hudi ... 6215
Données d'Upsert ... 6220
Supprimer un enregistrement ... 6221
Lire à partir d'un jeu de données Hudi ... 6223
Utiliser la CLI Hudi ... 6225
... 6226
Historique des sorties de Hudi ... 6226
Hue ... 6230
Versions de Hue ... 6230
Fonctionnalités prises en charge et non prises en charge de Hue sur Amazon EMR ... 6232
Considérations ... 6233
Performances avec de grandes tables de métadonnées Hue ... 6234
Incompatibilité entre les versions de Hue ... 6235
Connexion à l'interface utilisateur Web de Hue ... 6235
Utilisation de Hue avec une base de données distante dans Amazon RDS ... 6236
Résolution des problèmes ... 6238
Configurations avancées pour Hue ... 6239
Configuration de Hue pour les utilisateurs LDAP ... 6239
Historique des versions de Hue ... 6242

Iceberg 6308
Comment fonctionne Iceberg 6309
Utiliser un cluster Iceberg 6311
Utiliser un cluster Iceberg avec Spark 6311
Utiliser un cluster Iceberg avec Trino 6320
Utiliser un cluster Iceberg avec Flink 6322
Utiliser un cluster Iceberg avec Hive 6327
Historique des versions d'Iceberg 6332
Notes de mise à jour d'Iceberg par version 6333
Bloc-notes Jupyter 6336
EMR Studio 6336
Bloc-notes EMR 6336
JupyterHub 6337
Création d'un cluster avec JupyterHub 6341
Éléments à prendre en compte lors de l'utilisation JupyterHub sur Amazon EMR 6343
Configuration JupyterHub 6344
Configuration de la persistance pour les blocs-notes dans Amazon S3 6345
Connexion au nœud principal et aux serveurs de bloc-notes 6346
JupyterHub configuration et administration 6347
Ajout d'utilisateurs et d'administrateurs de bloc-notes Jupyter 6349
Installation de noyaux et de bibliothèques supplémentaires 6361
JupyterHub historique des publications 6366
Livy 6406
Activation de HTTPS 6408
Historique des versions de Livy 6409
MXNet 6457
MXNet historique des versions 6459
Oozie 6488
Utilisation d'Oozie avec une base de données distante dans Amazon RDS 6491
Configurer la version Java pour Oozie 6493
Historique des versions d'Oozie 6494
Notes de mise à jour d'Oozie par version 6555
Phoenix 6558
Création d'un cluster avec Phoenix 6561
Personnalisation des configurations Phoenix 6562
Clients Phoenix 6563

Historique des sorties de Phoenix 6566
Pig 6648
Soumission de travail Pig 6650
Soumettre le travail de Pig à l'aide de la console Amazon EMR 6651
Soumission d'un travail Pig à l'aide d' AWS CLI 6652
Appel de fonctions définies par l'utilisateur à partir de Pig 6653
Appel de fichiers JAR à partir de Pig 6654
Appel de scripts Python/Jython à partir de Pig 6654
Historique de versions Pig 6655
Presto 6724
Utilisation de Presto avec le catalogue de données AWS Glue 6726
Spécifier le catalogue de données AWS Glue comme métastore 6727
Autorisations IAM 6011
Considérations relatives à l'utilisation du catalogue de données AWS Glue 6732
Utilisation de S3 Select Pushdown 6733
S3 Select Pushdown est-il adapté à mon application ? 6733
Considérations et restrictions 6734
Activation de S3 Select Pushdown avec PrestoDB ou Trino 6734
Ajout de connecteurs de base de données 6735
Utilisation de SSL/TLS et de LDAPS 6736
Utilisation de l'authentification LDAP 6737
Activation du mode strict Presto 6745
Considérations 6747
Gestion de la perte d'instances Spot dans Presto 6748
Utilisation du dimensionnement automatique de Presto avec désaffectation gracieuse 6750
Considérations relatives à Presto sur Amazon EMR 6751
Exécutable en ligne de commande Presto 6751
Propriétés de déploiement Presto non configurables 6752
Installation de Prestodb 6752
Configuration d'EMRFS et de PrestOS3 FileSystem 6753
Paramètre par défaut de l'emprunt d'identité de l'utilisateur final 6754
Port par défaut de l'interface Web Presto 6754
Problème d'exécution du compartiment Hive dans certaines versions 6754
Historique des versions de Presto 6755
Spark 6813
Création d’un cluster Spark 6817

Exécution d'applications Spark avec Docker sur Amazon EMR 6.x 6820
Considérations 6820
Création d'une image Docker 6821
Utilisation d'images Docker depuis Amazon ECR 6822
Utiliser le catalogue AWS Glue Data Catalog avec Spark sur Amazon EMR 6828
Spécifier le catalogue de données AWS Glue comme métastore Apache Hive 6829
Spécifier le catalogue de données AWS Glue en tant que catalogue Apache Iceberg 6831
Autorisations IAM 6011
Considérations 6013
Utilisation d'une hiérarchie de plusieurs catalogues dans AWS Glue Data Catalog 6834
Comment les ressources du catalogue sont organisées 6834
Configuration d'un catalogue de données à utiliser avec Amazon EMR 6835
Autorisations d'accès aux ressources dans AWS Glue Data Catalog 6835
Configurer Spark pour accéder à une hiérarchie de catalogues multiples dans AWS Glue Data Catalog 6836
Considérations et limites relatives à une configuration à catalogues multiples 6839
Configuration de Spark 6839
Paramètres par défaut de Spark 6840
Configuration du récupérateur de mémoire 6842
`maximizeResourceAllocation` 6843
Comportement de mise hors service 6845
Variable d' ThriftServer environnement Spark 6848
Modification des paramètres Spark par défaut 6849
Migration d'Apache Log4j 1.x vers Log4j 2.x 6851
Optimisation des performances de Spark 6852
Exécution de requêtes adaptative 6852
Nettoyage dynamique de partition 6854
Aplatissement des sous-requêtes scalaires 6857
DISTINCT avant INTERSECT 6858
Jonction de filtre Bloom 6859
Réorganisation optimisée des jonctions 6859
Mise en cache des fragments de résultats 6860
Activation de la mise en cache des fragments de résultats Spark 6861
Considérations 6861
Utilisation de l'accélérateur RAPIDS 6864
Choix des types d'instance 6864

Configuration des applications 6864
Ajout d’une action d’amorçage 6872
Lancement de votre cluster 6873
Accès au shell de Spark 6873
Utiliser Amazon SageMaker Spark pour le machine learning 6875
Ecriture d'une application Spark 6875
Scala 6875
Java 6876
Python 6877
Amélioration des performances de Spark avec S3 6879
Utiliser S3 Select 6879
Valideur EMRFS optimisé pour S3 6884
Utiliser le protocole de validation optimisé pour EMRFS S3 6891
Nouvelle tentative pour les demandes S3 6898
Ajout d'une étape Spark 6901
Remplacement des paramètres de configuration par défaut de Spark 6904
Afficher l'historique de l'application Spark 6905
Accédez au site Web de Spark UIs 6905
Utilisation de Spark avec Amazon Kinesis Data Streams 6906
Utilisation de Spark sur Amazon Redshift 6906
Lancement d'une application Spark 6907
Authentification dans Amazon Redshift 6908
Lecture et écriture vers Amazon Redshift 6910
Considérations 6912
Historique des versions de Spark 6913
Sqoop 6980
Versions de Sqoop 6980
Éléments à prendre en compte avec Sqoop sur Amazon EMR 6982
Utilisation de Sqoop avec intégration HCatalog 6982
Prise en charge des bases de données et JDBC par Sqoop 6983
Sécurisation de votre mot de passe 6984
Historique des versions de Sqoop 6985
TensorFlow 7040
TensorFlow compilations par type d' EC2 instance Amazon 7042
Sécurité 7042
En utilisant TensorBoard 7042

TensorFlow historique des versions 7043
Tez 7068
Création d'un cluster avec Tez 7070
Configuration de Tez 7071
Exemple de configuration 7071
Ouverture fractionnée asynchrone Tez 7072
Interface utilisateur web Tez 7075
Serveur de chronologie 7076
Historique des versions de Tez 7076
Notes de publication de Tez par version 7122
Trino 7133
Histoire et design de Trino 7134
Concepts architecturaux 7135
Premiers pas avec Trino 7136
Suivez les étapes préalables à l'utilisation d'Amazon EMR avec Trino 7136
Lancement d'un cluster Amazon EMR avec Trino 7137
Connexion au nœud primaire 7139
Configuration de Trino 7140
Configuration des connecteurs pour Trino 7141
Surveillance 7141
Bonnes pratiques pour Trino sur Amazon EMR 7142
Principaux domaines d'intérêt pour l'amélioration des performances 7142
Collecter et utiliser les statistiques des tables 7144
Défis courants liés à l'évolutivité des charges de travail de Trino 7145
Considérations relatives à Trino 7150
Propriétés de déploiement Trino non configurables 7150
Considérations supplémentaires 7151
Historique des versions de Trino 7151
Notes de mise à jour de Trino par version 7152
Amazon EMR 7.6.0 7152
Amazon EMR 7.3.0 - Notes de mise à jour de Trino 7152
Amazon EMR 6.9.0 - Notes de mise à jour de Trino 7153
Zeppelin 7154
Éléments à prendre en compte lors de l'utilisation de Zeppelin sur Amazon EMR 7156
Historique des versions de Zeppelin 7157
ZooKeeper 7219

ZooKeeper historique des versions 7221
ZooKeeper notes de mise à jour par version 7261
Amazon EMR 7.8.0 - notes de mise à jour ZooKeeper 7262
Amazon EMR 7.5.0 : notes de mise à jour ZooKeeper 7262
Amazon EMR 7.4.0 : notes de mise à jour ZooKeeper 7262
Connecteurs et utilitaires 7263
Exportation, interrogation et jointure de tables dans DynamoDB 7263
Configuration d'une table Hive pour exécuter des commandes Hive 7265
Exemples de commandes Hive pour l'exportation, l'importation et l'interrogation de données 7274
Optimisation des performances 7284
Kinesis 7288
Que puis-je faire avec l'intégration d'Amazon EMR et Amazon Kinesis ? 7288
Analyse de points de contrôle des flux Amazon Kinesis 7288
Considérations sur les performances 7290
Planifiez une analyse Amazon Kinesis avec Amazon EMR 7290
Migration du connecteur Spark Kinesis vers le SDK 2.x pour Amazon EMR 7.0 7290
S3 DistCp (s3-dist-cp) 7300
DistCp Options S3 7301
Ajouter S3 DistCp en tant qu'étape dans un cluster 7309
Nettoyage après l'échec de DistCp tâches S3 7311
Exécuter des commandes et des scripts sur un cluster 7313
Soumettre une étape JAR personnalisée pour exécuter un script ou une commande 7314
Autres moyens d'utiliser `command-runner.jar` 7315
AWS Glossaire 7317
..... 7318

À propos des versions Amazon EMR

Une version Amazon EMR est un ensemble d'applications open-source issues de l'écosystème big data. Chaque version comprend plusieurs applications, composants et fonctionnalités Big Data que vous sélectionnez pour installer et configurer Amazon EMR lorsque vous créez un cluster. Les applications sont empaquetées à l'aide d'un système basé sur Apache BigTop, un projet open source associé à l'écosystème Hadoop. Ce guide fournit des informations sur les applications incluses dans les versions Amazon EMR.

Pour plus de plus amples informations sur la prise en main et l'utilisation d'Amazon EMR, consultez la section Guide de gestion Amazon EMR.

Lorsque vous lancez un cluster, vous pouvez choisir parmi plusieurs versions d'Amazon EMR. Vous pouvez ainsi tester et utiliser des versions d'applications qui répondent à vos exigences de compatibilité. Vous spécifiez le numéro de version avec l'étiquette de version. Les étiquettes de version sont sous la forme `emr-x.x.x`. Par exemple, `emr-7.9.0`.

À partir d'Amazon EMR 5.18.0, vous pouvez utiliser le référentiel d'artefacts d'Amazon EMR pour générer le code de votre tâche en fonction des versions exactes des bibliothèques et des dépendances qui sont disponibles avec des versions spécifiques d'Amazon EMR. Pour de plus amples informations, veuillez consulter Vérification des dépendances à l'aide du référentiel d'artefacts d'Amazon EMR.

Abonnez-vous au flux RSS des notes de mise à jour d'Amazon EMR pour recevoir des mises à jour lorsqu'une nouvelle version d'Amazon EMR est disponible.

Détails de la dernière version, tels que les versions de l'application, les notes de mise à jour, les composants et les classifications de configuration d'Amazon EMR série 7.x, 6.x et 5.x :

- Amazon EMR version 7.9.0
- Amazon EMR version 6.15.0
- Amazon EMR version 5.36.2

> **Note**
>
> Les nouvelles versions d'Amazon EMR sont mises à disposition dans différentes régions sur une période de plusieurs jours, en commençant par la première région à la date de

sortie initiale. Il est possible que la dernière version ne soit pas disponible dans votre région pendant cette période.

Notes de mise à jour pour les dernières versions d'Amazon EMR et historique de toutes les versions :

- Nouveautés
- Archive des notes de publication Amazon EMR

Historique complet des versions des applications dans chaque version d'Amazon EMR :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Détails de chaque version d'Amazon EMR et différences entre les versions d'une même série, le cas échéant :

- Versions Amazon EMR 7.x
- Versions Amazon EMR 6.x
- Versions Amazon EMR 5.x
- Versions Amazon EMR 4.x
- Versions AMI 2.x et 3.x d'Amazon EMR

Support standard d'Amazon EMR

Comprendre les versions d'Amazon EMR

La version Amazon EMR fournit les logiciels nécessaires pour exécuter vos applications open source sur EC2 des plateformes EKS ou sans serveur. Une version d'Amazon EMR est composée de trois éléments : l'environnement d'exécution, les moteurs principaux et les extras.

- Environnement d'exécution : l'environnement d'exécution inclut le système d'exploitation inclus dans Amazon Machine Image (AMI) ou toute autre image de conteneur lorsque vous lancez votre

cluster ou conteneur (par exemple, Amazon Linux 2023). Il inclut également les environnements d'exécution de langage tels que le JDK Amazon Corretto, ainsi que d'autres outils connexes installés par l'image.

- Moteurs principaux : il s'agit notamment des principaux logiciels open source tels qu'Apache Hive et des formats de table ouverts tels qu'Apache Iceberg.
- Extras — Il s'agit notamment de bibliothèques pratiques et de packages Python tels que `mariadb-connector-java` des logiciels open source tels qu'Apache Pig.

La version d'Amazon EMR que vous sélectionnez regroupe tous ces composants dans une nouvelle version d'Amazon Machine Images (AMI) ou d'images de conteneur basées sur Amazon Linux (AL). Amazon EMR gère la même version AL pour tous les nœuds d'un cluster, d'un conteneur EKS ou d'une application sans serveur. Amazon EMR vise à obtenir le dernier environnement d'exécution dans les 90 jours suivant leur publication par Amazon Linux, et les nouvelles versions open source de Core Engines dans les 90 jours suivant leur sortie en amont. Les extras sont publiés selon les besoins. Vous trouverez la liste des logiciels dans les notes de mise à jour de chaque version d'Amazon EMR.

Versionnage des versions

Pour vous aider à comprendre l'ampleur des modifications apportées à chaque version, Amazon EMR utilise le versionnement sémantique. Le versionnement sémantique suit un format de. `<major>.<minor>.<patch>` À titre d'exemple, considérons la version 7.0.0.

Le premier chiffre indique la version majeure, ou version majeure. Les versions majeures apportent généralement des modifications, des améliorations et de nouvelles fonctionnalités substantielles à l'environnement d'exécution ou au moteur principal qui ne sont pas rétrocompatibles. Par exemple, la version majeure 7.x utilise Amazon Linux 2023 avec le JDK Amazon Corretto 17 par défaut. Par conséquent, cette version contient plusieurs modifications majeures par rapport à la version majeure 6.x, qui est fournie sur Amazon Linux 2 et avec le JDK Amazon Corretto 8 par défaut.

Le deuxième chiffre indique la version mineure, ou version mineure. Les versions mineures sont des versions non rétrocompatibles qui contiennent des modifications, des améliorations et des fonctionnalités incrémentielles apportées aux moteurs principaux et aux extras, ainsi que de nouvelles fonctionnalités. Par exemple, la version 6.15 utilise Apache Spark 3.4.1, alors que la version 6.11.1 est fournie avec Apache Spark 3.3.2.

Le troisième chiffre indique une version de correctif, ou une version de correctif. Les versions de correctif sont destinées à être des versions rétrocompatibles contenant des correctifs et des mises

à jour pour Core Engine et les options supplémentaires, mais pas de nouvelles fonctionnalités ni de mises à niveau de version OSS. Par exemple, les versions 6.11.1 et 6.11.0 contiennent toutes deux la même version OSS d'Apache Spark 3.3.2. Depuis la version 6.6.x, la dernière version de correctif propose le dernier correctif d'environnement d'exécution disponible. L'environnement d'exécution n'affecte pas la publication du correctif d'Amazon EMR. Par exemple, 6.11.1 a été lancé avec la version 2.0.20240109.0 d'AL du 19 décembre 2023 au 10 janvier 2024, mais avec la version AL 2.0.20240124.0 du 11 janvier 2024. Certaines versions d'Amazon EMR on EKS vous permettent d'ajouter des balises ou des suffixes aux versions sémantiques afin de vous proposer des fonctionnalités alternatives. Par exemple, la `emr-6.15.0-latest` sortie d'Amazon EMR sur EKS lance une version compatible avec Amazon Corretto 8 JDK, tandis que la sortie d'Amazon EMR sur EKS vous permet de lancer une version compatible avec `emr-6.15.0-java17-latest` Amazon Corretto 17 JDK.

Politique de prise en charge

Introduction

Amazon EMR a pour objectif de publier de nouvelles versions mineures au moins une fois tous les 90 jours et de fournir un support pour les versions mineures pendant une période de 24 mois à compter de la date de sortie. Ce support couvre l'environnement d'exécution et les moteurs principaux avec leurs dépendances associées, mais ne couvre pas les extras ni leurs dépendances associées. L'environnement d'exécution inclut également des composants de APIs plate-forme propriétaires nécessaires au démarrage, à l'arrêt et au fonctionnement des clusters et des applications. Ce calendrier de publication cohérent garantit un cycle prévisible, ce qui vous permet de planifier, de tester et de passer plus facilement à une version offrant support et sécurité.

À quoi s'attendre avec le Support standard

Le support standard fournit des correctifs sur les tickets de support technique pour les problèmes que vous rencontrez avec les composants Runtime Environment et Core Engines dans le cadre des configurations recommandées. Tous les correctifs sont soumis à disponibilité. La configuration recommandée est l'utilisation d'Amazon EMR sans modification, ajout ou modification des fichiers binaires et des configurations présents dans une version d'Amazon EMR, à l'exception de ceux figurant dans la documentation Amazon EMR. Amazon EMR déploie des correctifs sur le dernier correctif, la version mineure ou majeure de la version d'Amazon EMR dans les 90 jours suivant la vérification du correctif par nos soins. Amazon EMR applique automatiquement des correctifs lorsque vous lancez un nouvel EMR sur un EC2 cluster, que vous lancez un nouveau conteneur Amazon EMR sur EKS ou que vous déclenchez une nouvelle tâche EMR sans serveur. Les composants

supplémentaires sont fournis pour des raisons de commodité et Amazon EMR ne fournit pas de correctifs liés aux suppléments.

Composants

Le support standard couvre les composants de l'environnement d'exécution et des moteurs principaux, par exemple le système d'exploitation, les environnements d'exécution des langages et les principaux logiciels open source tels qu'Apache Hive et Apache Iceberg. Vous trouverez la liste complète des composants pris en charge pour chaque version dans À propos des versions Amazon EMR Amazon Linux FAQs.

La liste suivante décrit le support que nous fournissons pour les différents types de composants dans le cadre du support standard :

- Composants de l'environnement d'exécution : les composants de l'environnement d'exécution recevront des correctifs sur les tickets de support technique. Les correctifs sont classés comme suit : (a) bogues critiques, (b) problèmes critiques de corruption de données et (c) problèmes de sécurité critiques. Si vous êtes éligible, Amazon EMR rétroportera les correctifs pour Runtime Environment vers les anciennes versions. Pour des raisons de compatibilité opérationnelle avec les composants open source du moteur principal, certains composants de l'environnement d'exécution doivent conserver des versions spécifiques afin de ne pas endommager les applications des clients. Pour ces composants, Amazon EMR dépend de l'open source en amont pour la disponibilité des correctifs. Lorsque des correctifs seront disponibles en open source, nous fournirons la dernière version stable dans les 90 jours suivant leur vérification par Amazon EMR.
- Composants de Core Engines : Core Engines fournit les dernières versions pour de nombreux projets open source, chacun contenant des centaines de bibliothèques de dépendances transitives. Bien que les communautés open source qui gèrent ces projets tentent fréquemment de résoudre les problèmes et de résoudre les vulnérabilités et expositions courantes (CVE) connues, les dernières versions peuvent toujours contenir des bogues connus et. CVEs Amazon EMR dépend de l'open source en amont pour la disponibilité des correctifs et fournira la dernière version stable dans le cadre des composants du moteur principal dans les 90 jours suivant leur vérification par Amazon EMR. Dans certains cas, Amazon EMR peut fournir un correctif pour un CVE dans l'un des composants du moteur principal qui doit être résolu avant l'open source en amont. Amazon EMR vous fournit également une assistance technique et des correctifs pour les fonctionnalités ajoutées en plus de l'open source dans les configurations recommandées. Nous ne rétroportons pas les correctifs pour les composants du moteur principal vers d'anciennes versions de correctifs ou de versions mineures.

- Composants supplémentaires : Amazon EMR ne prend pas en charge les composants supplémentaires. Les composants supplémentaires sont des projets open source fournis pour des raisons de commodité et Amazon EMR ne fournit pas de solutions aux problèmes rencontrés avec ces derniers. Toute demande d'assistance ou correction peut être traitée par le biais de la communauté open source supportant ces composants.

Support standard : cycle de vie

Ce qui suit décrit les étapes du cycle de vie du Support standard :

- Support standard : les versions d'Amazon EMR sont éligibles au support standard 24 mois à compter de la date de sortie. Vous pouvez créer des tickets de support technique et vous attendre à des mises à jour pour les problèmes que vous rencontrez avec ces versions.
- Fin du support : une fois le support standard terminé, les versions d'Amazon EMR entrent dans la phase de fin de support (eOS) pendant 12 mois. Les versions d'eOS ne sont pas éligibles au support technique et vous ne pourrez pas créer de tickets pour les clusters, les conteneurs ou les jobs exécutés sur ces versions. Les versions d'EoS ne recevront aucun correctif, correctif ou mise à jour. Les versions d'eOS seront supprimées de la console, mais continueront d'être disponibles via l'API et AWS CLI. Vous pouvez toujours continuer à exécuter des charges de travail sur les versions d'eOS. Nous vous recommandons vivement de migrer vers la dernière version d'Amazon EMR afin de continuer à recevoir les correctifs de sécurité, de rester éligible au support technique et de créer des tickets d'assistance en cas de besoin.
- Fin de vie : après la fin de la période de Support (EoS), les versions sont considérées comme étant en fin de vie (EoL). Bien que vous puissiez continuer à exécuter des clusters EoL, Amazon EMR se réserve le droit de supprimer les versions EoL de l'API et du SDK pour des raisons de case-by-case sécurité et d'exploitation. Nous vous recommandons vivement de migrer vers la dernière version d'Amazon EMR, car les versions d'EoL peuvent être supprimées de l'API et du SDK dans des cas exceptionnels.

Support de pont

Amazon EMR a annoncé cette nouvelle politique de support le 25 juillet 2024. En vertu de cette politique, les versions d'Amazon EMR publiées le 24 juillet 2022 ou avant cette date sont désormais désignées comme étant en fin de support. Toutefois, afin de vous donner plus de temps pour planifier et migrer vers des versions plus récentes, Amazon EMR proposera un support temporaire, équivalent au support standard, pour ces anciennes versions publiées dans les deux ans précédant cette annonce.

Après le 25 juillet 2024, vous pouvez consulter le statut actuel et les délais de support dans les notes de publication.

Le tableau suivant indique l'état de support de toutes les versions existantes d'Amazon EMR au moment de l'annonce de la politique, le 25 juillet 2024 :

Versions et périodes prises en charge

Version Amazon EMR	Date de sortie initiale	Date de fin du support standard	Date de début de fin du support	Date de début de fin de vie
7.2.0	25 juillet 2024	24 juillet 2026	25 juillet 2026	25 juillet 2027
7.1.0	23 avril 2024	22 avril 2026	23 avril 2026	23 avril 2027
7.0.0	19 décembre 2023	18 décembre 2025	19 décembre 2025	19 décembre 2026
5.36.x et 6.6.x — 6.15.x	9 mai 2022 au 13 novembre 2023	Support du pont jusqu'au 24 janvier 2026	25 janvier 2026	25 janvier 2026
• Série 6.x : 6.5.0 et versions antérieures • Série 5.x : 5.35.0 et versions inférieures • Séries 4.x, 3.x et 2.x	1er janvier 2013 au 30 mars 2022	Support du pont jusqu'au 24 juillet 2025	25 juillet 2025	25 juillet 2025

Considérations

Le support standard est disponible pour tous les modèles de déploiement d'Amazon EMR (EMR activé, EC2 Amazon EMR sur EKS et EMR sans serveur), dans toutes les régions où Amazon EMR est disponible, sans frais supplémentaires. Les clusters exécutés avec des configurations

recommandées sont automatiquement éligibles au support tel que décrit dans la politique. Vous n'avez donc pas besoin de prendre d'autres mesures pour activer le support.

- Le support standard prend uniquement en charge les composants requis pour les clusters Amazon EMR. Amazon EMR ne peut pas garantir les correctifs de sécurité et la disponibilité des correctifs dans le cas où les composants open source de Core Engine atteignent la fin de vie en amont, ou lorsque les mises à jour de sécurité ne sont plus disponibles pour les dépendances. Bien que vous puissiez choisir d'installer des options supplémentaires, Amazon EMR ne les prendra pas en charge ni ne prendra en charge leurs dépendances. Par exemple, vous pouvez installer des applications tierces dans votre AMI personnalisée pour renforcer la sécurité de votre cluster, installer des composants supplémentaires ou copier des objets à l'aide de scripts d'action bootstrap, ou via SSH dans votre cluster et mettre à niveau les versions de package par défaut. Amazon EMR ne prend pas en charge ces composants. Le support standard ne couvre pas les actions de démarrage, les packages, les bibliothèques fournis par le client, votre code personnalisé et les applications bring-your-own personnalisées que vous pouvez configurer Amazon EMR pour installer à votre convenance.
- Vos clusters existants ne seront pas affectés, quelle que soit la version d'Amazon EMR qu'ils utilisent. Vous pouvez continuer à exécuter les clusters existants sans interruption. Vous pouvez également continuer à lancer de nouveaux clusters et à exécuter des tâches sur toutes les versions existantes ou nouvelles. Toutes les versions existantes et les nouvelles versions au moment de l'entrée en vigueur de la politique sont couvertes par le Support Standard pendant 24 mois à compter de la date initiale de sortie de la version de lancement d'Amazon EMR. Amazon EMR vous fournira une assistance temporaire lors de l'annonce initiale de la politique. Pour bénéficier d'une assistance ininterrompue, nous vous conseillons de tester vos applications et de passer rapidement à la version la plus récente d'Amazon EMR.
- Amazon EMR ne modifiera pas les composants du Support standard sur les versions ou clusters existants. Amazon EMR se réserve toutefois le droit de respecter la fin de vie en amont sur une certaine case-by-case base et de supprimer ces composants dans les nouvelles versions, ou de passer la version existante au statut de fin de support (eOS) ou de fin de vie (eOL) dans des cas exceptionnels. Nous vous informerons de toute suppression par le biais des canaux disponibles.
- Lorsqu'une nouvelle région est lancée, Amazon EMR ne prend en charge que les versions d'Amazon EMR dans le cadre du support standard, publiées six mois avant la date de mise à disposition générale de la nouvelle région.
- Amazon EMR ne mettra pas automatiquement à jour vos clusters existants vers les dernières versions. Toutefois, vous pouvez choisir de mettre à jour les nouveaux clusters avec les dernières versions des correctifs si vous le souhaitez.

- Le coût de votre utilisation non prise en charge d'Amazon EMR sera pris en compte dans votre facture. AWS Même si vous utilisez Amazon EMR de manière non prise en charge, les coûts associés à cette utilisation font toujours partie de votre AWS consommation globale et seront inclus dans le calcul de vos frais de support.

Pour plus d'informations, contactez le Support aux AWS développeurs.

Versions Amazon EMR 7.x

Cette section contient les versions des applications, les notes de mise à jour, les versions des composants et les classifications de configuration disponibles dans chaque version 7.x d'Amazon EMR.

Lorsque vous lancez un cluster, vous pouvez choisir parmi plusieurs versions d'Amazon EMR. Vous pouvez ainsi tester et utiliser des versions d'applications qui répondent à vos exigences de compatibilité. Vous spécifiez le numéro de version avec l'étiquette de version. Les étiquettes de version sont sous la forme `emr-x.x.x`. Par exemple, `emr-7.9.0`.

Les nouvelles versions d'Amazon EMR sont mises à disposition dans différentes régions sur une période de plusieurs jours, en commençant par la première région à la date de sortie initiale. Il est possible que la dernière version ne soit pas disponible dans votre région pendant cette période.

Pour un tableau complet des versions des applications dans chaque version 7.x d'Amazon EMR, consultez Versions des applications dans les versions 7.x d'Amazon EMR.

Rubriques

- Versions des applications dans les versions 7.x d'Amazon EMR
- Amazon EMR version 7.9.0
- Amazon EMR version 7.8.0
- Amazon EMR version 7.7.0
- Amazon EMR version 7.6.0
- Amazon EMR version 7.5.0
- Amazon EMR version 7.4.0
- Amazon EMR version 7.3.0
- Amazon EMR version 7.2.0

- Amazon EMR version 7.1.0
- Amazon EMR version 7.0.0

Versions des applications dans les versions 7.x d'Amazon EMR

Pour obtenir un tableau complet répertoriant les versions d'applications disponibles dans chaque version d'Amazon EMR 7.x, ouvrez Versions des applications dans les versions 7.x d'Amazon EMR dans votre navigateur.

Amazon EMR version 7.9.0

Cycle de vie pris en charge en 7.9.0

Le tableau suivant décrit les dates de cycle de vie prises en charge pour la version 7.9.0 d'Amazon EMR.

Phase de support	Date
Date de sortie initiale	19 mai 2025
Support standard jusqu'à	19 mai 2027
Fin de l'assistance	20 mai 2027
Fin de vie	19 mai 2028

composants de la version 7.9.0

Amazon EMR 7.9.0 inclut les applications suivantes :

Moteurs principaux :

AmazonCloudWatchAgent, Delta, Flink, Hadoop,, Hive HBase, Hudi HCatalog, Livy, Iceberg, Phoenix, Spark, Tez, Trino JupyterEnterpriseGateway, gardien de zoo

Suppléments :

Hue JupyterHub, Oozie, Pig, Presto, Zeppelin TensorFlow

Pour plus d'informations sur les packages inclus, consultez la section Versions des applications dans les versions 7.x d'Amazon EMR. Pour plus d'informations sur les moteurs principaux et les options supplémentaires, consultez Amazon EMR Standard Support.

Versions de l'application 7.9.0

Cette version inclut les applications suivantes : AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, Oozie, PhoenixPig, Presto, Spark, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-7,9.0	emr-7,8.0	emr-7,7.0	emr-7.6.0
AWS SDK pour Java	2,31,16, 1,12,782	2,29,52, 1,12,780	2,29,44, 1,12,780	2,29,25, 1,12,779
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3,9, 3,11
Scala	2,1,18	2,1,18	2,1,18	2,1,18
AmazonCloudWatchAgent	1,300032,2	1,300032,2	1,300032,2	1,300032,2
Delta	3.3.0-amzn-1	3.3.0-amzn-0	3.2.1-amzn-2	3.2.1-amzn-1
Flink	1.20.0	1.20.0	1.20.0	1.20.0
HBase	2.6.2	2.6.1	2.6.1	2.6.1

	emr-7,9.0	emr-7,8.0	emr-7,7.0	emr-7.6.0
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.1	3.4.1	3.4.0	3.4.0
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0,15,0-amzn-6	0,15,0-amzn-5	0,15,0-amzn-4	0,15,0-amzn-3
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.7.1-amzn-2	1.7.1-amzn-1	1.7.1-amzn-0	1.6.1-amzn-2
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.1	5.2.1	5.2.1	5.2.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,287	0,287	0,287
Spark	3.5.5	3.5.4	3.5.3	3.5.3
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	467	467	467	457
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.3	3.9.3	3.9.2	3.9.2

notes de mise à jour de la 7.9.0

Les notes de mise à jour suivantes incluent des informations relatives à la version 7.9.0 d'Amazon EMR.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 7.9.0 application upgrades include AWS SDK for Java 2.31.16/1.12.782, Delta 3.3.0-amzn-1, Hadoop 3.4.1, Hudi 0.15.0-amzn-6, Iceberg 1.7.1-amzn-2, Spark 3.5.5, and Zookeeper 3.9.3.
- Support CM-CMK pour l'interface utilisateur persistante de l'application : cette fonctionnalité étend le chiffrement SSE appliqué aux journaux gérés dans le compartiment de service EMR afin de respecter la clé principale personnalisée gérée par le client (CM-CMK). Cette fonctionnalité permet aux clients de gérer l'accessibilité des journaux après la fin de leurs clusters.

Problèmes connus

- Lorsque l'option de journalisation pour EMR activé EC2 est désactivée, les journaux ne sont pas disponibles après la fermeture du cluster et l'application persistante UIs ne contient aucun journal.

Modifications, améliorations et problèmes résolus

- Obsolète d'Oozie — OSS Oozie étant passé en mode attique, la dernière version à inclure Apache Oozie est la version 7.9 d'Amazon EMR. Apache Oozie sera supprimé des versions 7.10 et supérieures d'Amazon EMR. Nous vous recommandons d'utiliser MWAA comme alternative à Oozie. Les clients peuvent se référer à l'utilitaire pour GitHub migrer oozie-to-airflow-emrles flux de travail Oozie existants vers MWAA.
- Emballage de conteneurs YARN — À partir de la version 7.9.0 d'Amazon EMR, la politique d'emballage de conteneurs est désormais disponible pour le planificateur de capacité YARN, qui repose sur la politique de placement de nœuds multiples de YARN.
- Le script hudi-cli.sh — Le script hudi-cli.sh est devenu obsolète dans les versions 7.9.0 et supérieures d'EMR. Les versions 7.9.0 et supérieures d'Amazon EMR utilisent le fichier .jar. hudi-cli-bundle
- Clusters EMR sur l'AMI AL2 023 compatible FIPS — À partir de la version 7.9 d'EMR, les clients peuvent lancer des clusters EMR sur l'AMI 023 compatible FIPS. AL2 Cette modification met à jour les signatures numériques et les algorithmes de hachage d'EMR afin de garantir que tous les packages EMR respectent les politiques cryptographiques recommandées pour la norme FIPS

140-3. Remarque : Cela ne couvre pas le statut de certification FIPS EMR AL2 ou 023. Reportez-vous AL2à la FAQ FIPS 023 pour obtenir les dernières informations sur le statut de validation de la norme FIPS 140-3.

- Latence par étapes — Cette version résout un problème selon lequel la latence des étapes augmentait lors de l'utilisation du chiffrement en transit sans sélectionner explicitement l'application Hadoop à installer.
- Lancement de clusters avec des volumes chiffrés EBS : cette version résout un problème selon lequel un appel KMS sur des volumes chiffrés par EBS pouvait entraîner l'échec du lancement d'un cluster EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250527.1	6.1.134-152.225 mai 2023	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7.2 250331.(	6.1.131-143.221 mai 2023	18 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre), Asie-Pacifique (Tokyo)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250303.(	6.1.129-138.220 août 2023	27 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250218,:	6.1.128-136.201. mai 2023	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250211.(	6.1.127-135.201. mai 2023	26 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

Versions Java 7.9.0 par défaut

Les versions 7.9 et supérieures d'Amazon EMR sont fournies par défaut avec Amazon Corretto 17 (basé sur OpenJDK) pour les applications compatibles avec Corretto 17 (JDK 17), à l'exception d'Apache Livy.

Le tableau suivant indique les versions Java par défaut pour les applications dans Amazon EMR 7.9.0. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
Zookeeper	17, 11, 8

Versions des composants 7.9.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous

nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
adot-java-agent	1,31,0	Agent Java qui collecte des métriques à partir des démons d'application.
delta	3.3.0-amzn-1	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-amazon-cloudwatch-agent	1,300032,2-amzn-0	Une application qui collecte des métriques internes au niveau du système et des métriques d'applica tion personnalisées à partir d'instances Amazon EC2 .
emr-ddb	5.7.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3,18,0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3,22,0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.18.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.39.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.18.0	Connecteur S3Select EMR
emr-wal-cli	1.3.4	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,70,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1,20,0-amzn-3	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1,20,0-amzn-3	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
hadoop-client	3.4.1-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.4.1-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.4.1-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.4.1-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	3.4.1-amzn-1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.4.1-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.4.1-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.4.1-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.4.1-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.4.1-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.4.1-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.6.2-amzn-0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	2.6.2-amzn-0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.6.2-amzn-0	HBase client en ligne de commande.
hbase-rest-server	2.6.2-amzn-0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.6.2-amzn-0	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.6.2-amzn-0	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-18	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-18	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-18	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-18	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-18	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	3.1.3-amzn-18	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-18	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0,15,0-amzn-6	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0,15,0-amzn-6	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0,15,0-amzn-6	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0,15,0-amzn-6	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1.7.1-amzn-2	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.

Composant	Version	Description
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.8.0 - incubation	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	12.3.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.2.1	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	6.0.0	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica

Composant	Version	Description
presto-coordinator	0,287-amzn-3	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0,287-amzn-3	Service pour exécuter les éléments d'une requête.
presto-client	0,287-amzn-3	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	467-amzn-2	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	467-amzn-2	Service pour exécuter les éléments d'une requête.
trino-client	467-amzn-2	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.3.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger

Composant	Version	Description
spark-client	3,5,5-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3,5,5-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3,5,5-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3,5,5-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	25.02.1-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
tensorflow	2.16.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-16	Les bibliothèques et l'applica tion tez YARN.
tez-on-worker	0.10.2-amzn-16	Les bibliothèques et l'applica tion tez YARN pour les nœuds de travail.
zeppelin-server	0.11.1	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.9.3-amzn-1	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.9.3-amzn-1	ZooKeeper client en ligne de commande.

7.9.0 classifications de configuration

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

classifications EMR-7.9.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j2	Modifiez les valeurs dans le fichier HBase hbase-log 4j2.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classifications	Description	Actions de reconfiguration
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.

Classifications	Description	Actions de reconfiguration
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.pr operties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-s ite.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.pr operties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.prope rties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_c onfig.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistan ce S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.

Classifications	Description	Actions de reconfiguration
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
phoenix-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.p roperties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Modifiez les valeurs dans ZooKeeper le fichier logback.xml.	Restarts Zookeeper server.
emr-metrics	Modifiez les paramètres métriques EMR pour ce nœud.	Restarts the CloudWatchAgent service.

journal des modifications de la version 7.9.0

Journal des modifications pour la version 7.9.0 et notes de version

Date	Événement	Description
19/05/2025	Publication de documents	Notes de mise à jour d'Amazon EMR 7.9.0 publiées pour la première fois

Date	Événement	Description
19/05/2025	Déploiement terminé	Amazon EMR 7.9.0 entièrement déployé dans toutes les régions prises en charge
19/05/2025	Version limitée initiale	Amazon EMR 7.9.0 a été déployé pour la première fois dans les régions commerciales initiales

Amazon EMR version 7.8.0

Cycle de vie pris en charge en 7.8.0

Le tableau suivant décrit les dates de cycle de vie prises en charge pour la version 7.8.0 d'Amazon EMR.

Phase de support	Date
Date de sortie initiale	07 mars 2025
Support standard jusqu'à	6 mars 2027
Fin de l'assistance	7 mars 2027
Fin de vie	7 mars 2028

composants de la version 7.8.0

Amazon EMR 7.8.0 inclut les applications suivantes :

Moteurs principaux :

AmazonCloudWatchAgent, Delta, Flink, Hadoop,, Hive HBase, Hudi HCatalog, Livy, Iceberg, Phoenix, Spark, Tez, Trino JupyterEnterpriseGateway, gardien de zoo

Suppléments :

Hue JupyterHub, Oozie, Pig, Presto, Zeppelin TensorFlow

Pour plus d'informations sur les packages inclus, consultez la section Versions des applications dans les versions 7.x d'Amazon EMR. Pour plus d'informations sur les moteurs principaux et les options supplémentaires, consultez Amazon EMR Standard Support.

Versions de l'application 7.8.0

Cette version inclut les applications suivantes : AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, Oozie, PhoenixPig, Presto, Spark, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-7,8.0	emr-7,7.0	emr-7.6.0	emr-7.5.0
AWS SDK pour Java	2,29,52, 1,12,780	2,29,44, 1,12,780	2,29,25, 1,12,779	2,28,8, 1,12,772
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3,9, 3,11
Scala	2,1,18	2,1,18	2,1,18	2,1,18
AmazonCloudWatchAgent	1,300032,2	1,300032,2	1,300032,2	1,300032,2
Delta	3.3.0-amzn-0	3.2.1-amzn-2	3.2.1-amzn-1	3.2.1-amzn-0
Flink	1.20.0	1.20.0	1.20.0	1.19.1

	emr-7,8.0	emr-7,7.0	emr-7.6.0	emr-7.5.0
Ganglia	-	-	-	-
HBase	2.6.1	2.6.1	2.6.1	2.5.10
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.1	3.4.0	3.4.0	3.4.0
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0,15,0-amzn-5	0,15,0-amzn-4	0,15,0-amzn-3	0,15,0-amzn-2
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.7.1-amzn-1	1.7.1-amzn-0	1.6.1-amzn-2	1.6.1-amzn-1
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	-
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.1	5.2.1	5.2.0	5.2.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,287	0,287	0,287
Spark	3.5.4	3.5.3	3.5.3	3.5.2
Sqoop	-	-	-	-
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1

	emr-7,8.0	emr-7,7.0	emr-7.6.0	emr-7.5.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	467	467	457	446
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.3	3.9.2	3.9.2	3.9.2

notes de mise à jour de la 7.8.0

Les notes de mise à jour suivantes incluent des informations relatives à la version 7.8.0 d'Amazon EMR.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 7.8.0 application upgrades include AWS SDK for Java 2.29.52/1.12.780, Delta 3.3.0-amzn-0, Hadoop 3.4.1, Hudi 0.15.0-amzn-5, Iceberg 1.7.1-amzn-1, Spark 3.5.4, and Zookeeper 3.9.3.

Problèmes connus

- Non applicable à cette version.

Modifications, améliorations et problèmes résolus

- Aucun changement majeur.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250527.1	6.1.134-152.225 mai 2023	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique (Taipei), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250512,(	6.1.134-152.225 mai 2023	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Bahreïn) (Émirats arabes unis), Canada (centre), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			, AWS GovCloud (États-Unis est), AWS GovCloud (États-Unis ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250428.1	6,1134-150,224 mai 2023	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Malaisie), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250414,(	6.1.132-147.221 mai 2023	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7.2 250331.(	6.1.131-143.221 mai 2023	18 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre), Asie-Pacifique (Tokyo)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250303.(	6.1.129-138.220 août 2023	27 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250218,:	6.1.128-136.201. mai 2023	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250211.(	6.1.127-135.201. mai 2023	26 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

Versions Java par défaut 7.8.0

Les versions 7.8 et supérieures d'Amazon EMR sont fournies par défaut avec Amazon Corretto 17 (basé sur OpenJDK) pour les applications compatibles avec Corretto 17 (JDK 17), à l'exception d'Apache Livy.

Le tableau suivant indique les versions Java par défaut pour les applications dans Amazon EMR 7.8.0. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
Zookeeper	17, 11, 8

Versions des composants 7.8.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous

nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
adot-java-agent	1,31,0	Agent Java qui collecte des métriques à partir des démons d'application.
delta	3.3.0-amzn-0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-amazon-cloudwatch-agent	1,300032,2-amzn-0	Une application qui collecte des métriques internes au niveau du système et des métriques d'applica tion personnalisées à partir d'instances Amazon EC2 .
emr-ddb	5.6.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3,17,0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3,21,0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.18.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.38.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.17.0	Connecteur S3Select EMR
emr-wal-cli	1.3.3	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,69,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1,20,0-amzn-2	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1,20,0-amzn-2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
hadoop-client	3.4.1-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.4.1-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.4.1-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.4.1-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	3.4.1-amzn-0	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.4.1-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.4.1-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.4.1-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.4.1-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.4.1-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.4.1-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.6.1-amzn-2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	2.6.1-amzn-2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.6.1-amzn-2	HBase client en ligne de commande.
hbase-rest-server	2.6.1-amzn-2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.6.1-amzn-2	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.6.1-amzn-2	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-17	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-17	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-17	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-17	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-17	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	3.1.3-amzn-17	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-17	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0,15,0-amzn-5	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0,15,0-amzn-5	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0,15,0-amzn-5	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0,15,0-amzn-5	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1.7.1-amzn-1	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.

Composant	Version	Description
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.8.0 - incubation	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	12.3.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.2.1	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	6.0.0	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica

Composant	Version	Description
presto-coordinator	0,287-amzn-2	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0,287-amzn-2	Service pour exécuter les éléments d'une requête.
presto-client	0,287-amzn-2	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	467-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	467-amzn-1	Service pour exécuter les éléments d'une requête.
trino-client	467-amzn-1	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.3.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger

Composant	Version	Description
spark-client	3,5,4-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3,5,4-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3,5,4-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3,5,4-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	24.12.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
tensorflow	2.16.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-15	Les bibliothèques et l'applica tion tez YARN.
tez-on-worker	0.10.2-amzn-15	Les bibliothèques et l'applica tion tez YARN pour les nœuds de travail.
zeppelin-server	0.11.1	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.9.3-amzn-0	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.9.3-amzn-0	ZooKeeper client en ligne de commande.

7.8.0 classifications de configuration

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

classifications EMR-7.8.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classifications	Description	Actions de reconfiguration
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.

Classifications	Description	Actions de reconfiguration
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.pr operties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-s ite.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.pr operties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.prope rties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_c onfig.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistan ce S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.

Classifications	Description	Actions de reconfiguration
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
phoenix-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.p roperties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Modifiez les valeurs dans ZooKeeper le fichier logback.xml.	Restarts Zookeeper server.
emr-metrics	Modifiez les paramètres métriques EMR pour ce nœud.	Restarts the CloudWatchAgent service.

journal des modifications de la version 7.8.0

Journal des modifications pour la version 7.8.0 et les notes de version

Date	Événement	Description
07/03/2025	Publication de documents	Notes de mise à jour d'Amazon EMR 7.8.0 publiées pour la première fois

Date	Événement	Description
06/03/2025	Déploiement terminé	Amazon EMR 7.8.0 entièrement déployé dans toutes les régions prises en charge
28/02/2025	Version limitée initiale	Amazon EMR 7.8.0 a été déployé pour la première fois dans les régions commerciales initiales

Amazon EMR version 7.7.0

Cycle de vie pris en charge en 7.7.0

Le tableau suivant décrit les dates de cycle de vie prises en charge pour la version 7.7.0 d'Amazon EMR.

Phase de support	Date
Date de sortie initiale	6 février 2025
Support standard jusqu'à	5 février 2027
Fin de l'assistance	6 février 2027
Fin de vie	6 février 2028

composants de la version 7.7.0

Amazon EMR 7.7.0 inclut les applications suivantes :

Moteurs principaux :

AmazonCloudWatchAgent, Delta, Flink, Hadoop,, Hive HBase, Hudi HCatalog, Livy, Iceberg, Phoenix, Spark, Tez, Trino JupyterEnterpriseGateway, gardien de zoo

Suppléments :

Hue JupyterHub, Oozie, Pig, Presto, Zeppelin TensorFlow

Pour plus d'informations sur les packages inclus, consultez la section Versions des applications dans les versions 7.x d'Amazon EMR. Pour plus d'informations sur les moteurs principaux et les options supplémentaires, consultez Amazon EMR Standard Support.

Versions de l'application 7.7.0

Cette version inclut les applications suivantes : AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, Oozie, PhoenixPig, Presto, Spark, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-7,7.0	emr-7.6.0	emr-7.5.0	emr-7.4.0
AWS SDK pour Java	2,29,44, 1,12,780	2,29,25, 1,12,779	2,28,8, 1,12,772	2,28,8, 1,12,772
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3,9, 3,11
Scala	2,1,18	2,1,18	2,1,18	2,1,18
AmazonCloudWatchAgent	1,300032,2	1,300032,2	1,300032,2	1,300032,2
Delta	3.2.1-amzn-2	3.2.1-amzn-1	3.2.1-amzn-0	3.2.0-amzn-1
Flink	1.20.0	1.20.0	1.19.1	1.19.1

	emr-7,7.0	emr-7.6.0	emr-7.5.0	emr-7.4.0
Ganglia	-	-	-	-
HBase	2.6.1	2.6.1	2.5.10	2.5.5
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.4.0	3.4.0	3.4.0
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0,15,0-amzn-4	0,15,0-amzn-3	0,15,0-amzn-2	0,15,0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.7.1-amzn-0	1.6.1-amzn-2	1.6.1-amzn-1	1.6.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	-
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.1	5.2.0	5.2.0	5.2.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,287	0,287	0,287
Spark	3.5.3	3.5.3	3.5.2	3.5.2
Sqoop	-	-	-	1.4.7
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1

	emr-7,7.0	emr-7.6.0	emr-7.5.0	emr-7.4.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	467	457	446	446
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.2	3.9.2	3.9.2	3.9.2

Détails des composants de la version 7.7.0

Consultez le tableau suivant pour plus d'informations sur les packages Extras dans Amazon EMR 7.7.0.

Application	Type de demande	Package	Type de package	Version du package
AmazonCloudWatchAgent	Principal	hadoop-hdfs-journalnode	Suppléments	3.4.0-amzn-3-1.amzn2023
AmazonCloudWatchAgent	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
AmazonCloudWatchAgent	Principal	s3-dist-cp	Suppléments	de 2 h 37 à 1 h du mois 2023
AmazonCloudWatchAgent	Principal	gardien de zoo	Suppléments	3.9.2. amzn.1-1. amzn2023
AmazonCloudWatchAgent	Principal	zookeeper-server	Suppléments	3.9.2. amzn.1-1. amzn2023
Flink	Principal	hadoop-client	Suppléments	3.4.0.amzn.3-1.amzn2023
Flink	Principal	hadoop-hdfs	Suppléments	3.4.0.amzn.3-1.amzn2023

Application	Type de demande	Package	Type de package	Version du package
Flink	Principal	hadoop-hdfs-datanode	Suppléments	3.4.0.amzn.3-1.amzn2023
Flink	Principal	hadoop-hdfs-journalnode	Suppléments	3.4.0.amzn.3-1.amzn2023
Flink	Principal	hadoop-hdfs-namenode	Suppléments	3.4.0.amzn.3-1.amzn2023
Flink	Principal	hadoop-httpfs	Suppléments	3.4.0.amzn.3-1.amzn2023
Flink	Principal	Hadoop KMS	Suppléments	3.4.0.amzn.3-1.amzn2023
Flink	Principal	hadoop-mapreduce	Suppléments	3.4.0.amzn.3-1.amzn2023
Flink	Principal	hadoop-mapreduce-historyserver	Suppléments	3.4.0.amzn.3-1.amzn2023
Flink	Principal	fil Hadoop	Suppléments	3.4.0.amzn.3-1.amzn2023
Flink	Principal	hadoop-yarn-nodemanager	Suppléments	3.4.0.amzn.3-1.amzn2023
Flink	Principal	hadoop-yarn-proxyserver	Suppléments	3.4.0.amzn.3-1.amzn2023
Flink	Principal	hadoop-yarn-resourcemanager	Suppléments	3.4.0.amzn.3-1.amzn2023

Application	Type de demande	Package	Type de package	Version du package
Flink	Principal	hadoop-yarn-timelineserver	Suppléments	3.4.0.amzn.3-1.amzn2023
Flink	Principal	hudi	Suppléments	0.15.0.amzn.4-1.amzn2023
Flink	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
Flink	Principal	hadoop	Suppléments	3.4.0.amzn.3-1.amzn2023
HBase	Principal	s3-dist-cp	Suppléments	de 2 h 37 à 1 h du mois 2023
HCatalog	Principal	hadoop-yarn-proxyserver	Suppléments	3.4.0.amzn.3-1.amzn2023
HCatalog	Principal	mariadb-connector-java	Suppléments	
Hive	Principal	hadoop-httpfs	Suppléments	3.4.0.amzn.3-1.amzn2023
Hive	Principal	hadoop-yarn-proxyserver	Suppléments	3.4.0.amzn.3-1.amzn2023
Hive	Principal	hive-hbase	Suppléments	3.1.3. amzn.16-1.amzn2023
Hive	Principal	mariadb-connector-java	Suppléments	2.7.2-1
Hive	Principal	s3-dist-cp	Suppléments	de 2 h 37 à 1 h du mois 2023

Application	Type de demande	Package	Type de package	Version du package
Phoenix	Principal	s3-dist-cp	Suppléments	de 2 h 37 à 1 h du mois 2023
Spark	Principal	noyau de données Spark	Suppléments	3.5.3.amzn.1-1.amzn2023
Tez	Principal	hadoop-mapreduce	Suppléments	3.4.0.amzn.3-1.amzn2023
Tez	Principal	hadoop-mapreduce-historyserver	Suppléments	3.4.0.amzn.3-1.amzn2023
Tez	Principal	hadoop-yarn-proxyserver	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	bigtop-utils	Suppléments	1.2.0-1.amzn2023
Trino	Principal	hadoop-client	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	hadoop-hdfs	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	hadoop-hdfs-datanode	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	hadoop-hdfs-journalnode	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	hadoop-hdfs-namenode	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	Hadoop KMS	Suppléments	3.4.0.amzn.3-1.amzn2023

Application	Type de demande	Package	Type de package	Version du package
Trino	Principal	hadoop-mapreduce	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	hadoop-mapreduce-historyserver	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	fil Hadoop	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	hadoop-yarn-nodemanager	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	hadoop-yarn-proxyserver	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	hadoop-yarn-resourcemanager	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	hadoop-yarn-timelineserver	Suppléments	3.4.0.amzn.3-1.amzn2023
Trino	Principal	hive	Suppléments	3.1.3. amzn.16-1.amzn2023
Trino	Principal	hive-hcatalog-server	Suppléments	3.1.3. amzn.16-1.amzn2023
Trino	Principal	mariadb-connector-java	Suppléments	2.7.2-1
Trino	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
Trino	Principal	gardien de zoo	Suppléments	3.9.2. amzn.1-1.amzn2023

Application	Type de demande	Package	Type de package	Version du package
Trino	Principal	zookeeper-server	Suppléments	3.9.2. amzn.1-1. amzn2023
Trino	Principal	bigtop-groovy	Suppléments	2.5.4-1. amzn2023
Trino	Principal	bigtop-jsvc	Suppléments	1.2.4-1. amzn2023
Trino	Principal	hadoop	Suppléments	3.4.0.amz n.3-1.amzn2023
Trino	Principal	hadoop-lzo	Suppléments	0.4.19-1 amzn2023
Trino	Principal	catalogue hive-h	Suppléments	3.1.3. amzn.16-1 .amzn2023
Trino	Principal	hive-jdbc	Suppléments	3.1.3. amzn.16-1 .amzn2023
Zookeeper	Principal	emrfs	Suppléments	3.9.2. amzn.1-1. amzn2023
Zookeeper	Principal	hadoop-client	Suppléments	3.9.2. amzn.1-1. amzn2023
Zookeeper	Principal	hadoop-hdfs	Suppléments	3.9.2. amzn.1-1. amzn2023
Zookeeper	Principal	hadoop-hdfs-datanode	Suppléments	3.9.2. amzn.1-1. amzn2023
Zookeeper	Principal	hadoop-hdfs-journalnode	Suppléments	3.9.2. amzn.1-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Zookeeper	Principal	hadoop-hdfs-namenode	Suppléments	3.9.2. amzn.1-1. amzn2023
Zookeeper	Principal	hadoop-httpfs	Suppléments	3.9.2. amzn.1-1. amzn2023
Zookeeper	Principal	Hadoop KMS	Suppléments	3.9.2. amzn.1-1. amzn2023
Zookeeper	Principal	hadoop-mapreduce	Suppléments	3.4.0.amzn.3-1.amzn2023
Zookeeper	Principal	hadoop-mapreduce-historyserver	Suppléments	3.4.0.amzn.3-1.amzn2023
Zookeeper	Principal	fil Hadoop	Suppléments	3.4.0.amzn.3-1.amzn2023
Zookeeper	Principal	hadoop-yarn-nodemanager	Suppléments	3.4.0.amzn.3-1.amzn2023
Zookeeper	Principal	hadoop-yarn-proxyserver	Suppléments	3.4.0.amzn.3-1.amzn2023
Zookeeper	Principal	hadoop-yarn-resourcemanager	Suppléments	3.4.0.amzn.3-1.amzn2023
Zookeeper	Principal	hadoop-yarn-timelineserver	Suppléments	3.4.0.amzn.3-1.amzn2023
Zookeeper	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023

notes de mise à jour de la 7.7.0

Les notes de mise à jour suivantes incluent des informations relatives à la version 7.7.0 d'Amazon EMR.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 7.7.0 application upgrades include Delta 3.2.1-amzn-2, Hudi 0.15.0-amzn-4, Iceberg 1.7.1-amzn-0, and Trino 467.

Problèmes connus

- Non applicable à cette version.

Modifications, améliorations et problèmes résolus

- La version d'Iceberg utilisée depuis EMR 7.7.0 ne prend plus en charge Java 8. En outre, Iceberg est exclu des images Java 8 suivantes : `emr-7.7.0-java8-latest` et`emr-7.7.0-spark-rapids-java8-latest`.
- Lake Formation pour le FGAC avec Amazon EMR sur EKS — Avec les versions 7.7 et supérieures d'Amazon EMR, vous pouvez tirer parti de Lake AWS Formation pour appliquer des contrôles d'accès précis aux tables du catalogue de données soutenues par Amazon S3. Cette fonctionnalité vous permet de configurer des contrôles d'accès au niveau des tables, des lignes, des colonnes et des cellules pour les requêtes de lecture dans votre Amazon EMR on EKS Spark Jobs. Pour plus d'informations, consultez la section Utilisation d'Amazon Amazon EMR sur EKS avec AWS Lake Formation pour un contrôle d'accès précis.

-

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250527.1	6.1.134-152.225 mai 2023	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique (Taipei), Canada (centre),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250512,(	6.1.134-152.225 mai 2023	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Bahreïn) (Émirats arabes unis), Canada (centre), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			, AWS GovCloud (États-Unis est), AWS GovCloud (États-Unis ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250428.1	6,1134-150,224 mai 2023	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourne), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Malaisie), Asie-Pacifique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250414,(	6.1.132-147.221 mai 2023	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7.2 250331.(	6.1.131-143.221 mai 2023	18 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre), Asie-Pacifique (Tokyo)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250303.(	6.1.129-138.220 août 2023	27 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250218,:	6.1.128-136.201. mai 2023	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250211.(	6.1.127-135.201. mai 2023	26 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

Versions Java 7.7.0 par défaut

Les versions 7.7 et supérieures d'Amazon EMR sont fournies par défaut avec Amazon Corretto 17 (basé sur OpenJDK) pour les applications compatibles avec Corretto 17 (JDK 17), à l'exception d'Apache Livy.

Le tableau suivant indique les versions Java par défaut pour les applications dans Amazon EMR 7.7.0. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
Zookeeper	17, 11, 8

Versions des composants 7.7.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous

nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
adot-java-agent	1,31,0	Agent Java qui collecte des métriques à partir des démons d'application.
delta	3.2.1-amzn-2	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-amazon-cloudwatch-agent	1,300032,2-amzn-0	Une application qui collecte des métriques internes au niveau du système et des métriques d'applica tion personnalisées à partir d'instances Amazon EC2 .
emr-ddb	5.6.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3,16,0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3,20,0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.18.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.37.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.16.0	Connecteur S3Select EMR
emr-wal-cli	1.3.2	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,68,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1,20,0-amzn-1	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1,20,0-amzn-1	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
hadoop-client	3.4.0-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.4.0-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.4.0-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.4.0-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	3.4.0-amzn-3	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.4.0-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.4.0-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.4.0-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.4.0-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.4.0-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.4.0-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.6.1-amzn-1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	2.6.1-amzn-1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.6.1-amzn-1	HBase client en ligne de commande.
hbase-rest-server	2.6.1-amzn-1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.6.1-amzn-1	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.6.1-amzn-1	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-16	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-16	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-16	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-16	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-16	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	3.1.3-amzn-16	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-16	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0,15,0-amzn-4	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0,15,0-amzn-4	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0,15,0-amzn-4	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0,15,0-amzn-4	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1.7.1-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.

Composant	Version	Description
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.8.0 - incubation	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	12.3.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.2.1	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	6.0.0	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica

Composant	Version	Description
presto-coordinator	0,287-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0,287-amzn-1	Service pour exécuter les éléments d'une requête.
presto-client	0,287-amzn-1	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	467-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	467-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	467-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.3.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger

Composant	Version	Description
spark-client	3.5.3-amzn-1	Clients de ligne de commande Spark.
spark-history-server	3.5.3-amzn-1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.5.3-amzn-1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.5.3-amzn-1	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	24.10.1-amzn-1	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
tensorflow	2.16.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-14	Les bibliothèques et l'applica tion tez YARN.
tez-on-worker	0.10.2-amzn-14	Les bibliothèques et l'applica tion tez YARN pour les nœuds de travail.
zeppelin-server	0.11.1	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.9.2-amzn-1	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.9.2-amzn-1	ZooKeeper client en ligne de commande.

7.7.0 classifications de configuration

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

classifications EMR-7.7.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classifications	Description	Actions de reconfiguration
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.prop erties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.

Classifications	Description	Actions de reconfiguration
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.pr operties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-s ite.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.pr operties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.prope rties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_c onfig.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistan ce S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.

Classifications	Description	Actions de reconfiguration
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
phoenix-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.p roperties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Modifiez les valeurs dans ZooKeeper le fichier logback.xml.	Restarts Zookeeper server.
emr-metrics	Modifiez les paramètres métriques EMR pour ce nœud.	Restarts the CloudWatchAgent service.

journal des modifications de la version 7.7.0

Journal des modifications pour la version 7.7.0 et notes de version

Date	Événement	Description
07/02/2025	Publication de documents	Notes de mise à jour d'Amazon EMR 7.7.0 publiées pour la première fois

Date	Événement	Description
06/02/2025	Déploiement terminé	Amazon EMR 7.7.0 entièrement déployé dans toutes les régions prises en charge
30-01-2025	Version limitée initiale	Amazon EMR 7.7.0 a été déployé pour la première fois dans les régions commerciales initiales

Amazon EMR version 7.6.0

Cycle de vie pris en charge en 7.6.0

Le tableau suivant décrit les dates de cycle de vie prises en charge pour la version 7.6.0 d'Amazon EMR.

Phase de support	Date
Date de sortie initiale	10 janvier 2025
Support standard jusqu'à	09 janvier 2027
Fin de l'assistance	10 janvier 2027
Fin de vie	10 janvier 2028

composants de la version 7.6.0

Amazon EMR 7.6.0 inclut les applications suivantes :

Moteurs principaux :

AmazonCloudWatchAgent, Delta, Flink, Hadoop,, Hive HBase, Hudi HCatalog, Livy, Iceberg, Phoenix, Spark, Tez, Trino JupyterEnterpriseGateway, gardien de zoo

Suppléments :

Hue JupyterHub, Oozie, Pig, Presto, Zeppelin TensorFlow

Pour plus d'informations sur les packages inclus, consultez la section Versions des applications dans les versions 7.x d'Amazon EMR. Pour plus d'informations sur les moteurs principaux et les options supplémentaires, consultez Amazon EMR Standard Support.

Versions de l'application 7.6.0

Cette version inclut les applications suivantes : AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, Oozie, PhoenixPig, Presto, Spark, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-7.6.0	emr-7.5.0	emr-7.4.0	emr-7.3.0
AWS SDK pour Java	2,29,25, 1,12,779	2,28,8, 1,12,772	2,28,8, 1,12,772	2,25,70, 1,12,747
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3,9, 3,11
Scala	2,1,18	2,1,18	2,1,18	2,1,18
AmazonCloudWatchAgent	1,300032,2	1,300032,2	1,300032,2	1,300032,2
Delta	3.2.1-amzn-1	3.2.1-amzn-0	3.2.0-amzn-1	3.2.0-amzn-0
Flink	1.20.0	1.19.1	1.19.1	1.18.1

	emr-7.6.0	emr-7.5.0	emr-7.4.0	emr-7.3.0
Ganglia	-	-	-	-
HBase	2.6.1	2.5.10	2.5.5	2.4,17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.4.0	3.4.0	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0,15,0-amzn-3	0,15,0-amzn-2	0,15,0-amzn-1	0,15,0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.6.1-amzn-2	1.6.1-amzn-1	1.6.1-amzn-0	1.5.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	-
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.0	5.2.0	5.2.0	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,287	0,287	0,285
Spark	3.5.3	3.5.2	3.5.2	3.5.1
Sqoop	-	-	1.4.7	1.4.7
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1

	emr-7.6.0	emr-7.5.0	emr-7.4.0	emr-7.3.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	457	446	446	442
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.2	3.9.2	3.9.2	3.9.1

7.6.0 vulnérabilités et risques courants

Le tableau suivant répertorie tout CVEs ce qui n'a aucun impact sur les clusters EMR qui s'exécutent sur les configurations Amazon EMR 7.6.0 recommandées. Amazon EMR dépend de l'open source en amont pour la disponibilité des correctifs et fournira la dernière version stable dans le cadre des composants du moteur principal dans les 90 jours suivant la vérification des correctifs par Amazon EMR.

Ce tableau a été mis à jour le 10 janvier 2025.

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-15250	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-15250
CVE-2020-17521	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-17521
CVE-2024-12798	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-12798
CVE-2024-12801	BAS	https://nvd.nist. gov/vuln/detail/CVE-2024-12801
CVE-2024-47554	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-47554

ID CVE	Sévérité	URL des détails du CVE
CVE-2_1471	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_1471
CVE-2_25857	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25857
CVE-2_38749	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_38749
CVE-2_38750	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_38750
CVE-2_38751	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_38751
CVE-2_38752	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_38752
CVE-2_41854	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_41854
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-avril 2004
CVE-2023-6378	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2024-36114	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-36114
CVE-2015-1832	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2015-1832

ID CVE	Sévérité	URL des détails du CVE
CVE-2018-1313	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2018-1313
CVE-2019-10219	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2019-10219
CVE-2020-10693	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-10693
CVE-2020-13949	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-13955	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-13955
CVE-2020-29582	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-29582
CVE-2020-8908	BAS	https://nvd.nist. gov/vuln/detail/CVE-2020-8908
CVE-2021-28170	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-28170
CVE-2021-34538	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-38153	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-38153
CVE-2_24329	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_24329
CVE-2_25647	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25647
CVE-2_36364	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_36364

ID CVE	Sévérité	URL des détails du CVE
CVE-2_39135	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_39135
CVE-2_42889	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_42889
CVE-2_46337	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_46337
CVE-2_46751	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_46751
CVE-2023-1370	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-20863	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-20863
CVE-2023-20883	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-20883
CVE-2023-25194	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-33201	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-33201
CVE-2023-33202	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-33202
CVE-2023-34453	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34454

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-34455	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-3635	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-39410	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-42503	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-42503
CVE-2023-43642	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-5072	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2023-51074	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-51074
CVE-2024-23945	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-23945
CVE-2024-25710	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-25710
CVE-2024-26308	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-26308
CVE-2024-29857	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-29857
CVE-2024-30171	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-30171
CVE-2024-31141	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-31141

ID CVE	Sévérité	URL des détails du CVE
CVE-2024-38808	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-38808
CVE-2024-38820	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-38820
CVE-2024-47561	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-47561
CVE-2024-52046	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-52046
CVE-2024-53990	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-53990
CVE-2024-56128	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-56128
CVE-2024-7254	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-7254
GHSA-CHFM-68VV-PVW5	BAS	Aucun lien n'est disponible.
CVE-2024-47535	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-47535
CVE-2012-5783	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2012-5783
CVE-2012-6153	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2012-6153
CVE-2016-5018	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2016-6796	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-6796

ID CVE	Sévérité	URL des détails du CVE
CVE-2018-10237	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2018-10237
CVE-2018-11798	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2018-11798
CVE-2019-0205	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-02-05
CVE-2020-15522	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-15522
CVE-2020-1945	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-1945
CVE-2021-36373	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-36373
CVE-2021-36374	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-36374
CVE-2018-1000180	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-1000180
CVE-2020-26939	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-26939
CVE-2020-28052	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-28052
CVE-2021-29425	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-29425
CVE-2015-5237	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2017-15095	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-15095

ID CVE	Sévérité	URL des détails du CVE
CVE-2017-17485	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2017-7525	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-7525
CVE-2018-11307	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14719
CVE-2018-14720	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19362
CVE-2018-5968	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-5968

ID CVE	Sévérité	URL des détails du CVE
CVE-2018-7489	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-7489
CVE-2019-12086	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2019-12384	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2019-12384
CVE-2019-12814	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2019-12814
CVE-2019-14379	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2019-17267	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17267

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-17531	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17531
CVE-2019-20330	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-14060	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14060

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-14061	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-35490	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36183

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-36184	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2021-22569	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-22569

ID CVE	Sévérité	URL des détails du CVE
CVE-2021-22570	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-22570
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2023-50387	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-50387
CVE-2023-50868	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-50868
CVE-2023-52428	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-52428
CVE-2024-23454	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-23454
CVE-2024-23944	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-23944
CVE-2024-25638	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-25638
CVE-2024-29025	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-29025
CVE-2024-29131	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-29131
CVE-2024-29133	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-29133

ID CVE	Sévérité	URL des détails du CVE
CVE-2024-51504	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-51504
CVE-2024-6763	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-6763
CVE-2024-8184	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-8184
CVE-2_40149	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40149
CVE-2_40150	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40150
CVE-2_45685	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45685
CVE-2_45693	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45693
CVE-2023-1436	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2013-7285	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2013-7285
CVE-2014-8125	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2014-8125
CVE-2016-3674	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-3674
CVE-2017-7957	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-7957
CVE-2019-10173	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-10173

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-26217	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-26217
CVE-2020-26258	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-26258
CVE-2021-21341	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-21341
CVE-2021-21342	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-21342
CVE-2021-21343	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-21343
CVE-2021-21344	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-21344
CVE-2021-21345	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-21345
CVE-2021-21346	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-21346
CVE-2021-21347	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-21347
CVE-2021-21348	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-21348
CVE-2021-21349	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-21349
CVE-2021-21350	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-21350
CVE-2021-21351	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-21351

ID CVE	Sévérité	URL des détails du CVE
CVE-2021-29505	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-29505
CVE-2021-39139	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39139
CVE-2021-39141	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39141
CVE-2021-39144	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39144
CVE-2021-39145	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39145
CVE-2021-39146	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39146
CVE-2021-39147	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39147
CVE-2021-39148	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39148
CVE-2021-39149	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39149
CVE-2021-39150	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39150
CVE-2021-39151	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39151
CVE-2021-39152	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39152
CVE-2021-39153	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39153

ID CVE	Sévérité	URL des détails du CVE
CVE-2021-39154	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-39154
CVE-2021-41411	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-41411
CVE-2021-43859	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-43859
CVE-2_1415	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_1415
CVE-2_40151	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40151
CVE-2_41966	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4/1966
CVE-2024-47072	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-47072
CVE-2019-10172	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-10172
CVE-2019-10202	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-10202
CVE-2_45688	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45688
CVE-2023-0833	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-0833
CVE-2024-21634	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2024-9823	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-9823

ID CVE	Sévérité	URL des détails du CVE
CVE-2_37866	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_37866
CVE-2012-0881	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2012-0881
CVE-2_23437	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_23437
CVE-2020-13936	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2_37865	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_37865
CVE-2009-2625	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2009-2625
CVE-2017-10355	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2017-10355
CVE-2021-31684	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2020-14338	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-14338
CVE-2018-2799	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2018-2799
CVE-2021-27568	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-27568
CVE-2013-4002	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2020-13956	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-13956

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-25649	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2023-26049	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-26049
CVE-2024-43382	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-43382
CVE-2023-34462	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-34462
CVE-2021-21290	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-21290
CVE-2023-44487	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-44487
CVE-2019-20445	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20445
GHSA-XPW8-RCWV-8F8P	ÉLEVÉ	Aucun lien n'est disponible.
CVE-2021-37136	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2021-21295	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-21295
GHSA-58QW-P7QM-5RVH	BAS	Aucun lien n'est disponible.
CVE-2021-37533	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-37533
CVE-2019-20444	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20444

ID CVE	Sévérité	URL des détails du CVE
CVE-2_40152	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40152

Détails des composants de la version 7.6.0

Consultez le tableau suivant pour plus d'informations sur les packages Extras dans Amazon EMR 7.6.0.

Application	Type de demande	Package	Type de package	Version du package
AmazonCloudWatchAgent	Principal	hadoop-hdfs-journalnode	Suppléments	3.4.0.amzn.2-1.amzn2023
AmazonCloudWatchAgent	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
AmazonCloudWatchAgent	Principal	s3-dist-cp	Suppléments	2.36.0-1 amzn2023
AmazonCloudWatchAgent	Principal	gardien de zoo	Suppléments	3.9.2. amzn.0-1. amzn2023
AmazonCloudWatchAgent	Principal	zookeeper-server	Suppléments	3.9.2. amzn.0-1. amzn2023
Flink	Principal	hadoop-client	Suppléments	3.4.0.amzn.2-1.amzn2023
Flink	Principal	hadoop-hdfs	Suppléments	3.4.0.amzn.2-1.amzn2023
Flink	Principal	hadoop-hdfs-datanode	Suppléments	3.4.0.amzn.2-1.amzn2023

Application	Type de demande	Package	Type de package	Version du package
Flink	Principal	hadoop-hdfs-journalnode	Suppléments	3.4.0.amzn.2-1.amzn2023
Flink	Principal	hadoop-hdfs-namenode	Suppléments	3.4.0.amzn.2-1.amzn2023
Flink	Principal	hadoop-httpfs	Suppléments	3.4.0.amzn.2-1.amzn2023
Flink	Principal	Hadoop KMS	Suppléments	3.4.0.amzn.2-1.amzn2023
Flink	Principal	hadoop-mapreduce	Suppléments	3.4.0.amzn.2-1.amzn2023
Flink	Principal	hadoop-mapreduce-historyserver	Suppléments	3.4.0.amzn.2-1.amzn2023
Flink	Principal	fil Hadoop	Suppléments	3.4.0.amzn.2-1.amzn2023
Flink	Principal	hadoop-yarn-nodemanager	Suppléments	3.4.0.amzn.2-1.amzn2023
Flink	Principal	hadoop-yarn-proxyserver	Suppléments	3.4.0.amzn.2-1.amzn2023
Flink	Principal	hadoop-yarn-resourcemanager	Suppléments	3.4.0.amzn.2-1.amzn2023
Flink	Principal	hadoop-yarn-timelineserver	Suppléments	3.4.0.amzn.2-1.amzn2023

Application	Type de demande	Package	Type de package	Version du package
Flink	Principal	hudi	Suppléments	0.15.0.amzn.3-1.amzn2023
Flink	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
Flink	Principal	hadoop	Suppléments	3.4.0.amzn.2-1.amzn2023
HBase	Principal	s3-dist-cp	Suppléments	2.36.0-1 amzn2023
HCatalog	Principal	hadoop-yarn-proxyserver	Suppléments	3.4.0.amzn.2-1.amzn2023
HCatalog	Principal	mariadb-connector-java	Suppléments	
Hive	Principal	hadoop-httpfs	Suppléments	3.4.0.amzn.2-1.amzn2023
Hive	Principal	hadoop-yarn-proxyserver	Suppléments	3.4.0.amzn.2-1.amzn2023
Hive	Principal	hive-hbase	Suppléments	3.1.3. amzn.15-1. amzn2023
Hive	Principal	mariadb-connector-java	Suppléments	2.7.2-1
Hive	Principal	s3-dist-cp	Suppléments	2.36.0-1 amzn2023
Phoenix	Principal	s3-dist-cp	Suppléments	2.36.0-1 amzn2023

Application	Type de demande	Package	Type de package	Version du package
Spark	Principal	noyau de données Spark	Suppléments	3.5.1. amzn.0-1. amzn2023
Tez	Principal	hadoop-ma preduce	Suppléments	3.4.0.amz n.2-1.amzn2023
Tez	Principal	hadoop-ma preduce-h istoryserver	Suppléments	3.4.0.amz n.2-1.amzn2023
Tez	Principal	hadoop-yarn-proxyserver	Suppléments	3.4.0.amz n.2-1.amzn2023
Trino	Principal	bigtop-utils	Suppléments	1.2.0-1. amzn2023
Trino	Principal	hadoop-client	Suppléments	3.4.0.amz n.2-1.amzn2023
Trino	Principal	hadoop-hdfs	Suppléments	3.4.0.amz n.2-1.amzn2023
Trino	Principal	hadoop-hdfs-datanode	Suppléments	3.4.0.amz n.2-1.amzn2023
Trino	Principal	hadoop-hdfs-journalnode	Suppléments	3.4.0.amz n.2-1.amzn2023
Trino	Principal	hadoop-hdfs-namenode	Suppléments	3.4.0.amz n.2-1.amzn2023
Trino	Principal	Hadoop KMS	Suppléments	3.4.0.amz n.2-1.amzn2023
Trino	Principal	hadoop-ma preduce	Suppléments	3.4.0.amz n.2-1.amzn2023

Application	Type de demande	Package	Type de package	Version du package
Trino	Principal	hadoop-mapreduce-historyserver	Suppléments	3.4.0.amzn.2-1.amzn2023
Trino	Principal	fil Hadoop	Suppléments	3.4.0.amzn.2-1.amzn2023
Trino	Principal	hadoop-yarn-nodemanager	Suppléments	3.4.0.amzn.2-1.amzn2023
Trino	Principal	hadoop-yarn-proxyserver	Suppléments	3.4.0.amzn.2-1.amzn2023
Trino	Principal	hadoop-yarn-resourcemanager	Suppléments	3.4.0.amzn.2-1.amzn2023
Trino	Principal	hadoop-yarn-timelineserver	Suppléments	3.4.0.amzn.2-1.amzn2023
Trino	Principal	hive	Suppléments	3.1.3. amzn.12-1. amzn2023
Trino	Principal	hive-hcatalog-server	Suppléments	3.1.3. amzn.12-1. amzn2023
Trino	Principal	mariadb-connector-java	Suppléments	2.7.2-1
Trino	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
Trino	Principal	gardien de zoo	Suppléments	3.9.2. amzn.0-1. amzn2023
Trino	Principal	zookeeper-server	Suppléments	3.9.2. amzn.0-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Trino	Principal	bigtop-groovy	Suppléments	2.5.4-1.amzn2023
Trino	Principal	bigtop-jsvc	Suppléments	1.2.4-1.amzn2023
Trino	Principal	hadoop	Suppléments	3.4.0.amzn.2-1.amzn2023
Trino	Principal	hadoop-lzo	Suppléments	0.4.19-1amzn2023
Trino	Principal	catalogue hive-h	Suppléments	3.1.3. amzn.15-1. amzn2023
Trino	Principal	hive-jdbc	Suppléments	3.1.3. amzn.12-1. amzn2023
Zookeeper	Principal	emrfs	Suppléments	3.9.2. amzn.0-1.amzn2023
Zookeeper	Principal	hadoop-client	Suppléments	3.9.2. amzn.0-1.amzn2023
Zookeeper	Principal	hadoop-hdfs	Suppléments	3.9.2. amzn.0-1.amzn2023
Zookeeper	Principal	hadoop-hdfs-datanode	Suppléments	3.9.2. amzn.0-1.amzn2023
Zookeeper	Principal	hadoop-hdfs-journalnode	Suppléments	3.9.2. amzn.0-1.amzn2023
Zookeeper	Principal	hadoop-hdfs-namenode	Suppléments	3.9.2. amzn.0-1.amzn2023

Application	Type de demande	Package	Type de package	Version du package
Zookeeper	Principal	hadoop-httpfs	Suppléments	3.9.2. amzn.0-1. amzn2023
Zookeeper	Principal	Hadoop KMS	Suppléments	3.9.2. amzn.0-1. amzn2023
Zookeeper	Principal	hadoop-ma preduce	Suppléments	3.4.0.amz n.2-1.amzn2023
Zookeeper	Principal	hadoop-ma preduce-h istoryserver	Suppléments	3.4.0.amz n.2-1.amzn2023
Zookeeper	Principal	fil Hadoop	Suppléments	3.4.0.amz n.2-1.amzn2023
Zookeeper	Principal	hadoop-yarn-nodemanager	Suppléments	3.4.0.amz n.2-1.amzn2023
Zookeeper	Principal	hadoop-yarn-proxyserver	Suppléments	3.4.0.amz n.2-1.amzn2023
Zookeeper	Principal	hadoop-ya rn-resour cemanager	Suppléments	3.4.0.amz n.2-1.amzn2023
Zookeeper	Principal	hadoop-yarn-timelineserver	Suppléments	3.4.0.amz n.2-1.amzn2023
Zookeeper	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023

notes de publication de la version 7.6.0

Les notes de mise à jour suivantes incluent des informations relatives à la version 7.6.0 d'Amazon EMR.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 7.6.0 application upgrades include HBase 2.6.1, Flink 1.20.0, Spark 3.5.3, and Trino 457.

Problèmes connus

- Lorsque le dimensionnement géré est activé, le YARN ResourceManager (RM) rencontre un problème de blocage critique, le rendant insensible lors de la transition de l'état DECOMMISSIONING à l'état DECOMMISSIONED avec des opérations simultanées. Cela se produit en particulier dans les environnements de grands clusters. Lorsque cela se produit, les transitions entre l'état du nœud et l'état de l'application ne sont pas traitées, paralysant ainsi les fonctionnalités de base du. ResourceManager Pour contourner le problème, redémarrez la JVM RM à l'aide de la commande suivante sur le nœud principal :

  ```
  sudo systemctl restart hadoop-yarn-resourcemanager
  ```

- Lorsqu'un cluster EMR est démarré avec des planificateurs YARN autres que le `CapacityScheduler` ou `FairScheduler` (par exemple, le planificateur FIFO), la `YARNMemoryAvailablePercentage` métrique n'est pas poussée vers. CloudWatch Ce problème a un impact sur la réduction d'échelle dans les clusters compatibles avec le dimensionnement géré. Cela peut également affecter le dimensionnement géré personnalisé s'il dépend de la `YARNMemoryAvailablePercentage` métrique. Pour contourner le problème, utilisez le `CapacityScheduler` ou`FairScheduler`.

Modifications, améliorations et problèmes résolus

- CloudWatch l'agent est pris en charge sur les clusters de rôles d'exécution pour EMR 7.6 et versions ultérieures. Pour plus d'informations, consultez Amazon CloudWatch agent.
- À partir de la version 7.6.0 d'Amazon Elastic Map Reduce (EMR), le connecteur de système de fichiers S3A prend désormais en charge le chiffrement côté client Amazon S3.
- À partir de la version 7.6.0 d'EMR, le protocole TLS est pris en charge sur et sur les points de terminaison. HMaster RegionServer Pour plus d'informations, consultez le tableau suivant concernant le support TLS avec HBase : HBase.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250527.1	6.1.134-152.225 mai 2023	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique (Taipei), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250512,(	6.1.134-152.225 mai 2023	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Bahreïn) (Émirats arabes unis), Canada (centre), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			, AWS GovCloud (États-Unis est), AWS GovCloud (États-Unis ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250428.1	6,1134-150,224 mai 2023	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Malaisie), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250414,(	6.1.132-147.221 mai 2023	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7.2 250331.(	6.1.131-143.221 mai 2023	18 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre), Asie-Pacifique (Tokyo)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250303.(	6.1.129-138.220 août 2023	27 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250218,:	6.1.128-136.201. mai 2023	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250211.(	6.1.127-135.201. mai 2023	26 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250123.4	6.1.124-134.200. mai 2023	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250115,(	6.1.119-129.201.am zn2023	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5.2 240916	6 janvier 109 à 118,189 mai 2023	6 septembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

Versions Java 7.6.0 par défaut

Les versions 7.6 et supérieures d'Amazon EMR sont fournies par défaut avec Amazon Corretto 17 (basé sur OpenJDK) pour les applications compatibles avec Corretto 17 (JDK 17), à l'exception d'Apache Livy.

Le tableau suivant indique les versions Java par défaut pour les applications dans Amazon EMR 7.6.0. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
Zookeeper	17, 11, 8

Versions des composants 7.6.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
adot-java-agent	1,31,0	Agent Java qui collecte des métriques à partir des démons d'application.
delta	3.2.1-amzn-1	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-amazon-cloudwatch-agent	1,300032,2-amzn-0	Une application qui collecte des métriques internes au niveau du système et des métriques d'applica tion personnalisées à partir d'instances Amazon EC2 .
emr-ddb	5.6.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3,15,0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3,19,0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.18.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.36.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.15.0	Connecteur S3Select EMR
emr-wal-cli	1.3.1	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,67,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1,20,0-amzn-0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1,20,0-amzn-0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
hadoop-client	3.4.0-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.4.0-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.4.0-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.4.0-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	3.4.0-amzn-2	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.4.0-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.4.0-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.4.0-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.4.0-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.4.0-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.4.0-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.6.1-amzn-0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	2.6.1-amzn-0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.6.1-amzn-0	HBase client en ligne de commande.
hbase-rest-server	2.6.1-amzn-0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.6.1-amzn-0	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.6.1-amzn-0	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-15	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-15	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-15	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-15	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-15	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	3.1.3-amzn-15	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-15	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0,15,0-amzn-3	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0,15,0-amzn-3	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0,15,0-amzn-3	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0,15,0-amzn-3	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1.6.1-amzn-2	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.

Composant	Version	Description
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.8.0 - incubation	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	12.3.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.2.0	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	6.0.0	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica

Composant	Version	Description
presto-coordinator	0,287-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0,287-amzn-1	Service pour exécuter les éléments d'une requête.
presto-client	0,287-amzn-1	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	457-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	457-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	457-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.3.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger

Composant	Version	Description
spark-client	3.5.3-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.5.3-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.5.3-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.5.3-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	24.10.1-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
tensorflow	2.16.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-13	Les bibliothèques et l'applica tion tez YARN.
tez-on-worker	0.10.2-amzn-13	Les bibliothèques et l'applica tion tez YARN pour les nœuds de travail.
zeppelin-server	0.11.1	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.9.2-amzn-0	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.9.2-amzn-0	ZooKeeper client en ligne de commande.

Classifications de configuration 7.6.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

classifications EMR-7.6.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classifications	Description	Actions de reconfiguration
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.

Classifications	Description	Actions de reconfiguration
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.pr operties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-s ite.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.pr operties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.prope rties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_c onfig.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistan ce S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.

Classifications	Description	Actions de reconfiguration
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
phoenix-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.p roperties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Modifiez les valeurs dans ZooKeeper le fichier logback.xml.	Restarts Zookeeper server.
emr-metrics	Modifiez les paramètres métriques EMR pour ce nœud.	Restarts the CloudWatchAgent service.

journal des modifications de la version 7.6.0

Journal des modifications pour la version 7.6.0 et notes de version

Date	Événement	Description
13/01/2025	Publication de documents	Notes de mise à jour d'Amazon EMR 7.6.0 publiées pour la première fois

Date	Événement	Description
10-01-2025	Déploiement terminé	Amazon EMR 7.6.0 entièrement déployé dans toutes les régions prises en charge
26/12/2025	Première version	Amazon EMR 7.6.0 a été déployé pour la première fois dans les régions commerciales initiales

Amazon EMR version 7.5.0

Cycle de vie pris en charge en 7.5.0

Le tableau suivant décrit les dates de cycle de vie prises en charge pour la version 7.5.0 d'Amazon EMR.

Phase de support	Date
Date de sortie initiale	21 novembre 2024
Support standard jusqu'à	20 novembre 2026
Fin de l'assistance	21 novembre 2026
Fin de vie	21 novembre 2027

Pour obtenir une liste des vulnérabilités et des risques courants critiques et élevés (CVEs) qui n'affectent pas les clusters EMR dans le cadre des logiciels et des configurations recommandés, voir 7.5.0 Vulnérabilités et expositions courantes connues pour les moteurs principaux.

composants de la version 7.5.0

Amazon EMR 7.5.0 inclut les applications suivantes :

Moteurs principaux :

AmazonCloudWatchAgent, Delta, Flink, Hadoop,, Hive HBase, Hudi HCatalog, Livy, Iceberg, Phoenix, Spark, Tez, Trino JupyterEnterpriseGateway, gardien de zoo

Suppléments :

Hue JupyterHub, Oozie, Pig, Presto, Zeppelin TensorFlow

Pour plus d'informations sur les packages inclus, consultez la section Versions des applications dans les versions 7.x d'Amazon EMR. Pour plus d'informations sur les moteurs principaux et les options supplémentaires, consultez Amazon EMR Standard Support.

Versions de l'application 7.5.0

Cette version inclut les applications suivantes : AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, Oozie, PhoenixPig, Presto, Spark, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-7.5.0	emr-7.4.0	emr-7.3.0	emr-7.2.0
AWS SDK pour Java	2,28,8, 1,12,772	2,28,8, 1,12,772	2,25,70, 1,12,747	2,23,18, 1,12,705
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3,9, 3,11
Scala	2,1,18	2,1,18	2,1,18	2,1,17

	emr-7.5.0	emr-7.4.0	emr-7.3.0	emr-7.2.0
AmazonCloudWatchAgent	1,300032,2	1,300032,2	1,300032,2	1,300032,2
Delta	3.2.1-amzn-0	3.2.0-amzn-1	3.2.0-amzn-0	3.1.0
Flink	1.19.1	1.19.1	1.18.1	1.18.1
Ganglia	-	-	-	-
HBase	2.5.10	2.5.5	2.4,17	2.4,17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.4.0	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0,15,0-amzn-2	0,15,0-amzn-1	0,15,0-amzn-0	0,14.1-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.6.1-amzn-1	1.6.1-amzn-0	1.5.2-amzn-0	1,5.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.0	5.2.0	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-7.5.0	emr-7.4.0	emr-7.3.0	emr-7.2.0
Presto	0,287	0,287	0,285	0,285
Spark	3.5.2	3.5.2	3.5.1	3.5.1
Sqoop	-	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.16.1	2.16.1	2.11.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	446	446	442	436
Zeppelin	0.11.1	0.11.1	0.11.1	0,1,1
ZooKeeper	3.9.2	3.9.2	3.9.1	3.9.1

7.5.0 vulnérabilités et risques courants

Le tableau suivant répertorie tout CVEs ce qui n'a aucun impact sur les clusters EMR qui s'exécutent sur les configurations Amazon EMR 7.5.0 recommandées. Amazon EMR dépend de l'open source en amont pour la disponibilité des correctifs et fournira la dernière version stable dans le cadre des composants du moteur principal dans les 90 jours suivant la vérification des correctifs par Amazon EMR.

Ce tableau a été mis à jour le 21 novembre 2024.

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-35728	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2021-37533	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37533
CVE-2018-14721	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-14721

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-26939	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-26939
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2024-47554	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-47554
CVE-2021-38153	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-38153
CVE-2020-36188	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-10673	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-8908	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-8908
CVE-2_3509	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_41966	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-4/1966
CVE-2023-3635	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2020-25649	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2017-17485	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2024-25710	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-25710

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-25194	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
GHSA-58QW-P7QM-5RVH	NON TRIÉ	https://nvd.nist. gov/vuln/detail/GHSA-58 qw-p7qm-5rvh
CVE-2023-51074	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-51074
CVE-2023-20883	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-20883
CVE-04-2004	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004
CVE-2024-47561	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-47561
CVE-2023-2976	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2_45685	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45685
CVE-2023-33202	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-33202
CVE-2019-14892	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2023-33201	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-33201
CVE-2023-1436	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2018-10237	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-10237

ID CVE	Sévérité	URL des détails du CVE
CVE-2015-5237	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-2024-36114	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-36114
CVE-2024-7254	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-7254
CVE-2020-14061	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2024-23944	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-23944
CVE-2020-10969	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2023-52428	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-52428
CVE-2024-29857	BAS	https://nvd.nist. gov/vuln/detail/CVE-2024-29857
CVE-2024-30171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-30171
CVE-2_41854	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_41854
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2018-7489	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-7489

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-35491	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-11111	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-15250	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-15250
CVE-2020-9548	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2024-51504	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-51504
CVE-2023-6378	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-34453	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2024-29025	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-29025
CVE-2024-36124	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-36124
CVE-2013-4002	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2019-14379	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2020-36186	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2_45688	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_45688

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-17521	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-17521
CVE-2024-8184	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-8184
CVE-2020-14195	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2023-40167	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-40167
CVE-2024-6763	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-6763
CVE-2_25647	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_25647
CVE-2018-14719	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2018-14719
CVE-2024-29131	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-29131
CVE-2020-9546	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2018-1000180	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-1000180
CVE-2023-39410	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2021-22570	BAS	https://nvd.nist. gov/vuln/detail/CVE-2021-22570
CVE-2_46751	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_46751

ID CVE	Sévérité	URL des détails du CVE
CVE-2021-36373	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-36373
CVE-2023-26048	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-26048
CVE-2024-30172	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-30172
CVE-2_36364	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_36364
CVE-2_46337	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_46337
CVE-2019-16869	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-16869
CVE-2_38751	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_38751
CVE-2_1415	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_1415
CVE-2021-22569	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-22569
CVE-2024-35255	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-35255
CVE-2021-28170	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-28170
CVE-2_40149	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40149
CVE-2023-50868	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-50868

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-34454	haut	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2_38750	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_38750
CVE-2020-14060	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2023-42503	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-42503
CVE-2024-25638	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-25638
CVE-2024-29133	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-29133
CVE-2_37865	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_37865
CVE-2018-19361	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2018-19362
CVE-2020-10650	modérée	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-35490	modérée	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2024-23454	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-23454
CVE-2023-43642	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-43642

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-24750	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-11112	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2019-14439	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2023-34455	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2024-26308	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-26308
CVE-2019-20445	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-20445
CVE-2023-0833	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-0833
CVE-2020-36185	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2023-6481	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2_38752	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_38752
CVE-2020-1945	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-1945
CVE-2020-13936	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2023-36478	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-36478

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-10968	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-28052	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-28052
CVE-2021-43859	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-43859
CVE-2021-27568	BAS	https://nvd.nist. gov/vuln/detail/CVE-2021-27568
CVE-2020-11113	BAS	https://nvd.nist. gov/vuln/detail/CVE-2020-11113
CVE-2_40150	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40150
CVE-2023-5072	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2018-11307	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2024-21634	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2020-36518	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-36180	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2021-34538	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-29425	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-29425

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-8840	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9547	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-36179	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2019-16943	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2016-5018	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2024-34447	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-34447
CVE-2024-39676	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-39676
CVE-2021-21290	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-21290
CVE-2019-12086	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2018-14718	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2020-36181	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2_39135	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2_39135
CVE-2019-16335	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-16335

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-36184	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2_45693	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45693
CVE-2020-14062	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-15522	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-15522
CVE-2019-14540	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2017-15095	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2018-5968	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2019-14893	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-12814	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-12814
CVE-2020-10672	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-24616	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-13949	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2023-50387	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-50387

ID CVE	Sévérité	URL des détails du CVE
CVE-2018-12022	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2020-36189	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2021-37137	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2019-12384	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-12384
CVE-2023-26049	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-26049
CVE-2_1471	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_1471
CVE-2023-1370	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2019-16942	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2_40152	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40152
CVE-2023-20863	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-20863
CVE-2018-11798	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-11798
CVE-2019-20330	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2019-17531	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-17531

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-11620	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2021-31684	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2020-11619	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2_40151	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40151
CVE-2019-17267	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2020-36182	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2016-6796	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-02-05
CVE-2018-10054	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-10054
CVE-2018-14720	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-2_24329	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2_24329
CVE-2020-36183	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2023-44981	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-44981

ID CVE	Sévérité	URL des détails du CVE
CVE-2_37866	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_37866
CVE-2015-1832	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2023-38647	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-38647
CVE-2018-1313	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-1313
CVE-2018-12023	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2021-20190	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2024-38808	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-38808
CVE-2019-20444	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-20444
CVE-2012-0881	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2012-0881
CVE-2020-36187	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-13956	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13956
CVE-2_38749	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_38749
CVE-2017-7525	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-7525

ID CVE	Sévérité	URL des détails du CVE
CVE-2018-19360	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2021-37136	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2_42889	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_42889
CVE-2021-36374	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-36374
CVE-2020-13955	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13955
CVE-2_25857	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_25857

Détails des composants de la version 7.5.0

Consultez le tableau suivant pour plus d'informations sur les packages Extras dans Amazon EMR 7.5.0.

Application	Type de demande	Package	Type de package	Version du package
AmazonCloudWatchAgent	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.5-1. amzn2023
AmazonCloudWatchAgent	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
AmazonCloudWatchAgent	Principal	s3-dist-cp	Suppléments	2 h 33 - 1 h 00

Application	Type de demande	Package	Type de package	Version du package
AmazonCloudWatchAgent	Principal	gardien de zoo	Suppléments	3.9.1. amzn.1-1. amzn2023
AmazonCloudWatchAgent	Principal	zookeeper-server	Suppléments	3.9.1. amzn.1-1. amzn2023
Flink	Principal	hadoop-client	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-hdfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-hdfs-datanode	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-hdfs-namenode	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-httpfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	Hadoop KMS	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-mapreduce	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-mapreduce-historyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	fil Hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Flink	Principal	hadoop-yarn-nodemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-ya rn-resour cemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-yarn-timelineserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hudi	Suppléments	0.15.0.am zn.0-1.am zn2023
Flink	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
Flink	Principal	hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
HBase	Principal	s3-dist-cp	Suppléments	2 h 33 - 1 h 00
HCatalog	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
HCatalog	Principal	mariadb-c onnector-java	Suppléments	
Hive	Principal	hadoop-httpfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Hive	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Hive	Principal	hive-hbase	Suppléments	3.1.3. amzn.12-1 . amzn2023
Hive	Principal	mariadb-c onnector-java	Suppléments	2.7.2-1
Hive	Principal	s3-dist-cp	Suppléments	2 h 33 - 1 h 00
Phoenix	Principal	s3-dist-cp	Suppléments	2 h 33 - 1 h 00
Spark	Principal	noyau de données Spark	Suppléments	3.5.1. amzn.0-1. amzn2023
Tez	Principal	hadoop-ma preduce	Suppléments	3.3.6. amzn.5-1. amzn2023
Tez	Principal	hadoop-ma preduce-h istoryserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Tez	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	bigtop-utils	Suppléments	1.2.0-1. amzn2023
Trino	Principal	hadoop-client	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-hdfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-hdfs-datanode	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.5-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Trino	Principal	hadoop-hdfs-namenode	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	Hadoop KMS	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-mapreduce	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-mapreduce-historyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	fil Hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-yarn-nodemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-yarn-resourcemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-yarn-timelineserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hive	Suppléments	3.1.3. amzn.12-1. amzn2023
Trino	Principal	hive-hcatalog-server	Suppléments	3.1.3. amzn.12-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Trino	Principal	mariadb-connector-java	Suppléments	2.7.2-1
Trino	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
Trino	Principal	gardien de zoo	Suppléments	3.9.1. amzn.1-1. amzn2023
Trino	Principal	zookeeper-server	Suppléments	3.9.1. amzn.1-1. amzn2023
Trino	Principal	bigtop-groovy	Suppléments	2.5.4-1. amzn2023
Trino	Principal	bigtop-jsvc	Suppléments	1.2.4-1. amzn2023
Trino	Principal	hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-lzo	Suppléments	0.4.19-1 amzn2023
Trino	Principal	catalogue hive-h	Suppléments	3.1.3. amzn.12-1. amzn2023
Trino	Principal	hive-jdbc	Suppléments	3.1.3. amzn.12-1. amzn2023
Zookeeper	Principal	emrfs	Suppléments	2.64.0-1 amzn2023
Zookeeper	Principal	hadoop-client	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-hdfs	Suppléments	3.3.6. amzn.5-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Zookeeper	Principal	hadoop-hdfs-datanode	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-hdfs-namenode	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-httpfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	Hadoop KMS	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-mapreduce	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-mapreduce-historyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	fil Hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-yarn-nodemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-yarn-resourcemanager	Suppléments	3.3.6. amzn.5-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Zookeeper	Principal	hadoop-yarn-timelineserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023

notes de mise à jour de la 7.5.0

Les notes de mise à jour suivantes incluent des informations relatives à la version 7.5.0 d'Amazon EMR.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 7.5.0 application upgrades include Delta 3.2.1-amzn-0, HBase 2.5.10, Hudi 0.15.0-amzn-2, and Iceberg 1.6.1-amzn-1.
- Les journaux d'écriture anticipée (WAL) EMR prennent en charge la réplication entre clusters. Pour plus d'informations, consultez la section Réplication entre clusters EMR WAL.

Problèmes connus

- Apache Sqoop a été supprimé de la version 7.5 et des versions ultérieures d'Amazon EMR.
- Lorsque le dimensionnement géré est activé, le YARN ResourceManager (RM) rencontre un problème de blocage critique, le rendant insensible lors de la transition de l'état DECOMMISSIONING à l'état DECOMMISSIONED avec des opérations simultanées. Cela se produit en particulier dans les environnements de grands clusters. Lorsque cela se produit, les transitions entre l'état du nœud et l'état de l'application ne sont pas traitées, paralysant ainsi les fonctionnalités de base du. ResourceManager Pour contourner le problème, redémarrez la JVM RM à l'aide de la commande suivante sur le nœud principal :

  ```
  sudo systemctl restart hadoop-yarn-resourcemanager
  ```

- Lorsqu'un cluster EMR est démarré avec des planificateurs YARN autres que le `CapacityScheduler` ou `FairScheduler` (par exemple, le planificateur FIFO), la `YARNMemoryAvailablePercentage` métrique n'est pas poussée vers. CloudWatch Ce problème a un impact sur la réduction d'échelle dans les clusters compatibles avec le

dimensionnement géré. Cela peut également affecter le dimensionnement géré personnalisé s'il dépend de la `YARNMemoryAvailablePercentage` métrique. Pour contourner le problème, utilisez le `CapacityScheduler` ou`FairScheduler`.

Modifications, améliorations et problèmes résolus

- Amazon EMR utilise différentes versions de la TensorFlow bibliothèque, en fonction des types d'instances que vous choisissez pour votre cluster. Amazon EMR prend également en charge les clusters dotés de types TensorFlow d'instances aarch64 (instances Graviton) à partir d'EMR-7.5.0 et versions ultérieures.

  Le tableau suivant répertorie les builds par type d'instance pour EMR-7.5.0 :

EC2 types d'instances	TensorFlow construire		
M5 et C5	Tensorflow 2.16.1 avec optimisation Intel MKL		
P2, P4D, P5, G4DN, G5, G6 et GR6	Tensorflow 2.16.1 avec CUDA 12.3, cuDNN 8.9.7.29		
P3, P3DN, G3 et G3S	Tensorflow 2.16.1 avec CUDA 12.3, cuDNN 8.9.7.29, NCCL 2.20.3-1 Nvidia NCCL n'est disponible que sur les instances P3. Contrat de licence utilisateur final (EULA) — En utilisant les composants Nvidia sur Amazon EMR, vous acceptez les termes et conditions décrits dans le CLUF du produit.		

EC2 types d'instances	TensorFlow construire		
Tous les autres, y compris Graviton (EMR-7.5.0+)	Tensorflow 2.16.1		

- Un cluster Amazon EMR de AWS Lake Formation avec FGAC (Fine-Grained Access Control) peut désormais prendre en charge jusqu'à 256 étapes simultanées avec la configuration hdfs-site.xml suivante :

```
[
  {
    "Classification": "hdfs-site",
    "Properties": {
      "dfs.datanode.handler.count": "40",
      "dfs.datanode.max.transfer.threads": "16384",
      "dfs.namenode.handler.count": "64",
      "dfs.replication": "1"
    }
  },
  {
    "Classification": "hdfs-env",
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "HADOOP_DATANODE_HEAPSIZE": "5120"
        }
      }
    ],
    "Properties": {}
  }
]
```

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250527.1	6.1.134-152.225 mai 2023	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique (Taipei), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250512,(	6.1.134-152.225 mai 2023	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Bahreïn) (Émirats arabes unis), Canada (centre), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			, AWS GovCloud (États-Unis est), AWS GovCloud (États-Unis ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250428.1	6,1134-150,224 mai 2023	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourne), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Malaisie), Asie-Pacifique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250414,(	6.1.132-147.221 mai 2023	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7.2 250331.(	6.1.131-143.221 mai 2023	18 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre), Asie-Pacifique (Tokyo)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250303.(	6.1.129-138.220 août 2023	27 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250218,:	6.1.128-136.201. mai 2023	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250211.(	6.1.127-135.201. mai 2023	26 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250123.4	6.1.124-134.200. mai 2023	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250115,(	6.1.119-129.201.am zn2023	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5.2 240916	6 janvier 109 à 118,189 mai 2023	6 septembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

Versions Java 7.5.0 par défaut

Les versions 7.5 et supérieures d'Amazon EMR sont fournies par défaut avec Amazon Corretto 17 (basé sur OpenJDK) pour les applications compatibles avec Corretto 17 (JDK 17), à l'exception d'Apache Livy.

Le tableau suivant indique les versions Java par défaut pour les applications dans Amazon EMR 7.5.0. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17 ANS
Zeppelin	17, 11, 8
Pig	8
Zookeeper	17, 11, 8

Versions des composants 7.5.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
adot-java-agent	1,31,0	Agent Java qui collecte des métriques à partir des démons d'application.
delta	3.2.1-amzn-0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-amazon-cloudwatch-agent	1,300032,2-amzn-0	Une application qui collecte des métriques internes au niveau du système et des métriques d'applica tion personnalisées à partir d'instances Amazon EC2 .
emr-ddb	5.6.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3,14,0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3,18,0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.16.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.35.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.14.0	Connecteur S3Select EMR
emr-wal-cli	1.3.1	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,66,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.19.1-amzn-1	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.19.1-amzn-1	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
hadoop-client	3.4.0-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.4.0-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.4.0-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.4.0-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	3.4.0-amzn-1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.4.0-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.4.0-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.4.0-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.4.0-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.4.0-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.4.0-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.5.10-amzn-0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	2.5.10-amzn-0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.5.10-amzn-0	HBase client en ligne de commande.
hbase-rest-server	2.5.10-amzn-0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.5.10-amzn-0	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.5.10-amzn-0	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-14	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-14	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-14	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-14	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-14	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	3.1.3-amzn-14	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-14	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0,15,0-amzn-2	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0,15,0-amzn-2	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0,15,0-amzn-2	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0,15,0-amzn-2	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
iceberg	1.6.1-amzn-1	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.

Composant	Version	Description
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.8.0 - incubation	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	12.3.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.2.0	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	6.0.0	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica

Composant	Version	Description
presto-coordinator	0,287-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0,287-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0,287-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	446-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	446-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	446-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.3.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger

Composant	Version	Description
spark-client	3.5.2-amzn-1	Clients de ligne de commande Spark.
spark-history-server	3.5.2-amzn-1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.5.2-amzn-1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.5.2-amzn-1	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	24.08.1-amzn-1	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
tensorflow	2.16.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-12	Les bibliothèques et l'applica tion tez YARN.
tez-on-worker	0.10.2-amzn-12	Les bibliothèques et l'applica tion tez YARN pour les nœuds de travail.
zeppelin-server	0.11.1	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.9.2-amzn-0	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.9.2-amzn-0	ZooKeeper client en ligne de commande.

7.5.0 classifications de configuration

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

classifications EMR-7.5.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classifications	Description	Actions de reconfiguration
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.prop erties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.

Classifications	Description	Actions de reconfiguration
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.pr operties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-s ite.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.pr operties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.prope rties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_c onfig.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistan ce S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.

Classifications	Description	Actions de reconfiguration
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
phoenix-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.p roperties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Modifiez les valeurs dans ZooKeeper le fichier logback.xml.	Restarts Zookeeper server.
emr-metrics	Modifiez les paramètres métriques EMR pour ce nœud.	Restarts the CloudWatchAgent service.

journal des modifications 7.5.0

Journal des modifications pour la version 7.5.0 et notes de version

Date	Événement	Description
21/11/2024	Publication de documents	Notes de mise à jour d'Amazon EMR 7.5.0 publiées pour la première fois

Date	Événement	Description
21/11/2024	Déploiement terminé	Amazon EMR 7.5.0 entièrement déployé dans toutes les régions prises en charge
13/11/2024	Première version	Amazon EMR 7.5.0 a été déployé pour la première fois dans les régions commerciales initiales

Amazon EMR version 7.4.0

Cycle de vie pris en charge en 7.4.0

Le tableau suivant décrit les dates de cycle de vie prises en charge pour la version 7.4.0 d'Amazon EMR.

Phase de support	Date
Date de sortie initiale	13 novembre 2024
Support standard jusqu'à	12 novembre 2026
Fin de l'assistance	13 novembre 2026
Fin de vie	13 novembre 2027

Pour obtenir une liste des vulnérabilités et des risques courants critiques et élevés (CVEs) qui n'affectent pas les clusters EMR dans le cadre des logiciels et des configurations recommandés, voir 7.4.0 Vulnérabilités et expositions courantes connues pour les moteurs principaux.

composants de la version 7.4.0

Amazon EMR 7.4.0 inclut les applications suivantes :

Moteurs principaux :

AmazonCloudWatchAgent, Delta, Flink, Hadoop,, Hive HBase, Hudi HCatalog, Livy, Iceberg, Phoenix, Spark, Tez, Trino JupyterEnterpriseGateway, gardien de zoo

Suppléments :

Hue JupyterHub, Oozie, Pig, Presto, Sqoop, Zeppelin TensorFlow

Pour plus d'informations sur les packages inclus, consultez la section Versions des applications dans les versions 7.x d'Amazon EMR. Pour plus d'informations sur les moteurs principaux et les options supplémentaires, consultez Amazon EMR Standard Support.

Versions de l'application 7.4.0

Cette version inclut les applications suivantes : AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, Oozie, PhoenixPig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-7.4.0	emr-7.3.0	emr-7.2.0	emr-7.1.0
AWS SDK pour Java	2,28,8, 1,12,772	2,25,70, 1,12,747	2,23,18, 1,12,705	2,23,18, 1,12,656
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3,9, 3,11
Scala	2,1,18	2,1,18	2,1,17	2,1,17

	emr-7.4.0	emr-7.3.0	emr-7.2.0	emr-7.1.0
AmazonCloudWatchAgent	1,300032,2	1,300032,2	1,300032,2	1,300032,2
Delta	3.2.0-amzn-1	3.2.0-amzn-0	3.1.0	3.0.0
Flink	1.19.1	1.18.1	1.18.1	1.18.1
Ganglia	-	-	-	-
HBase	2.5.5	2.4,17	2.4,17	2.4,17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0,15,0-amzn-1	0,15,0-amzn-0	0,14.1-amzn-1	0.14.1-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.6.1-amzn-0	1.5.2-amzn-0	1,5.0-amzn-0	1.4.3-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.0	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-7.4.0	emr-7.3.0	emr-7.2.0	emr-7.1.0
Presto	0,287	0,285	0,285	0,284
Spark	3.5.2	3.5.1	3.5.1	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.16.1	2.11.0	2.11.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	446	442	436	435
Zeppelin	0.11.1	0.11.1	0,1,1	0,1,1
ZooKeeper	3.9.2	3.9.1	3.9.1	3.9.1

7.4.0 vulnérabilités et expositions courantes

Le tableau suivant répertorie tout CVEs ce qui n'a aucun impact sur les clusters EMR qui s'exécutent sur les configurations Amazon EMR 7.4.0 recommandées. Amazon EMR dépend de l'open source en amont pour la disponibilité des correctifs et fournira la dernière version stable dans le cadre des composants du moteur principal dans les 90 jours suivant la vérification des correctifs par Amazon EMR.

Ce tableau a été mis à jour le 15 novembre 2024.

ID CVE	Sévérité	URL des détails du CVE
CVE-2021-37533	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37533
CVE-2018-14721	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2020-26939	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-26939

ID CVE	Sévérité	URL des détails du CVE
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2024-47554	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-47554
CVE-2021-38153	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-38153
CVE-2_41966	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-4/1966
CVE-2023-3635	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2020-25649	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2017-17485	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2024-25710	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-25710
CVE-2023-25194	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
GHSA-58QW-P7QM-5RVH	NON TRIÉ	https://nvd.nist. gov/vuln/detail/GHSA-58 qw-p7qm-5rvh
CVE-2023-51074	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-51074
CVE-2023-20883	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-20883
CVE-04-2004	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004

ID CVE	Sévérité	URL des détails du CVE
CVE-2024-47561	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-47561
CVE-2023-2976	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2_45685	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45685
CVE-2023-33202	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-33202
CVE-2019-14892	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2023-33201	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-33201
CVE-2023-1436	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2018-10237	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-10237
CVE-2015-5237	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-2024-36114	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-36114
CVE-2024-7254	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-7254
CVE-2020-14061	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14061

ID CVE	Sévérité	URL des détails du CVE
CVE-2024-23944	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-23944
CVE-2020-10969	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2023-52428	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-52428
CVE-2024-29857	BAS	https://nvd.nist. gov/vuln/detail/CVE-2024-29857
CVE-2024-30171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-30171
CVE-2_41854	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_41854
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2018-7489	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-7489
CVE-2020-35491	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-11111	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-15250	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-15250
CVE-2020-9548	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2024-51504	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-51504

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-6378	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-34453	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2024-29025	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-29025
CVE-2024-36124	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-36124
CVE-2013-4002	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2019-14379	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2020-36186	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2_45688	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_45688
CVE-2020-17521	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-17521
CVE-2024-8184	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-8184
CVE-2020-14195	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2023-40167	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-40167
CVE-2024-6763	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-6763

ID CVE	Sévérité	URL des détails du CVE
CVE-2_25647	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_25647
CVE-2018-14719	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2018-14719
CVE-2024-29131	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-29131
CVE-2020-9546	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2018-1000180	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-1000180
CVE-2023-39410	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2021-22570	BAS	https://nvd.nist. gov/vuln/detail/CVE-2021-22570
CVE-2_46751	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_46751
CVE-2021-36373	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2021-36373
CVE-2023-26048	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-26048
CVE-2024-30172	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-30172
CVE-2_36364	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_36364
CVE-2_46337	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_46337

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-16869	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-16869
CVE-2_38751	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_38751
CVE-2_1415	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_1415
CVE-2021-22569	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-22569
CVE-2024-35255	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-35255
CVE-2021-28170	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-28170
CVE-2_40149	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40149
CVE-2023-50868	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-50868
CVE-2_38750	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_38750
CVE-2020-14060	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2023-42503	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-42503
CVE-2024-25638	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-25638
CVE-2024-29133	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-29133

ID CVE	Sévérité	URL des détails du CVE
CVE-2_37865	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_37865
CVE-2018-19361	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2018-19362
CVE-2020-10650	modérée	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-35490	modérée	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2024-23454	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-23454
CVE-2023-43642	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2020-24750	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-11112	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2019-14439	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2023-34455	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2024-26308	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2024-26308
CVE-2019-20445	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-20445

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-0833	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-0833
CVE-2020-36185	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2023-6481	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2012-5783	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2012-5783
CVE-2_38752	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2_38752
CVE-2020-1945	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-1945
CVE-2020-13936	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2023-36478	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2020-10968	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-28052	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-28052
CVE-2021-43859	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-43859
CVE-2021-27568	BAS	https://nvd.nist. gov/vuln/detail/CVE-2021-27568
CVE-2020-11113	BAS	https://nvd.nist. gov/vuln/detail/CVE-2020-11113

ID CVE	Sévérité	URL des détails du CVE
CVE-2_40150	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40150
CVE-2023-5072	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2018-11307	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2019-16943	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2_39135	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_39135
CVE-2019-16335	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2017-7525	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-7525
CVE-2018-19360	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2020-35728	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2023-34454	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2024-21634	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2020-36518	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2012-6153	BAS	https://nvd.nist. gov/vuln/detail/CVE-2012-6153

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-36180	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2021-34538	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-29425	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-29425
CVE-2020-8840	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9547	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-36179	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2016-5018	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2024-34447	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-34447
CVE-2024-39676	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-39676
CVE-2021-21290	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-21290
CVE-2019-12086	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2018-14718	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2020-36181	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36181

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-36184	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2_45693	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45693
CVE-2020-14062	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-15522	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2020-15522
CVE-2019-14540	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2017-15095	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2018-5968	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2019-14893	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-12814	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-12814
CVE-2020-10672	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-24616	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-13949	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2023-50387	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-50387

ID CVE	Sévérité	URL des détails du CVE
CVE-2018-12022	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2020-36189	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2021-37137	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2019-12384	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-12384
CVE-2023-26049	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-26049
CVE-2_1471	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_1471
CVE-2023-1370	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2019-16942	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2_40152	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40152
CVE-2023-20863	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-20863
CVE-2018-11798	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-11798
CVE-2019-20330	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2019-17531	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-17531

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-11620	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2021-31684	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2020-11619	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2_40151	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40151
CVE-2019-17267	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2020-36182	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2016-6796	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-02-05
CVE-2018-10054	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-10054
CVE-2018-14720	MOYEN	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-2_24329	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2_24329
CVE-2020-36183	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2023-44981	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-44981

ID CVE	Sévérité	URL des détails du CVE
CVE-2_37866	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_37866
CVE-2015-1832	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2023-38647	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-38647
CVE-2018-1313	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-1313
CVE-2018-12023	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2021-20190	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2024-38808	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-38808
CVE-2019-20444	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-20444
GHSA-CHFM-68VV-PVW5	ÉLEVÉ	https://github.com/advisories/GHSA-CHFM-68VV-PVW5
CVE-2012-0881	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2012-0881
CVE-2020-36187	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-13956	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13956
CVE-2_38749	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_38749

ID CVE	Sévérité	URL des détails du CVE
CVE-2021-37136	NON TRIÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2_42889	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_42889
CVE-2021-36374	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-36374
CVE-2020-13955	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13955
CVE-2_25857	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_25857

Détails des composants de la version 7.4.0

Consultez le tableau suivant pour plus d'informations sur les packages Extras dans Amazon EMR 7.4.0.

Application	Type de demande	Package	Type de package	Version du package
AmazonCloudWatchAgent	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.5-1. amzn2023
AmazonCloudWatchAgent	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
AmazonCloudWatchAgent	Principal	s3-dist-cp	Suppléments	2 h 33 - 1 h 00
AmazonCloudWatchAgent	Principal	gardien de zoo	Suppléments	3.9.1. amzn.1-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
AmazonCloudWatchAgent	Principal	zookeeper-server	Suppléments	3.9.1. amzn.1-1. amzn2023
Flink	Principal	hadoop-client	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-hdfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-hdfs-datanode	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-hdfs-namenode	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-httpfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	Hadoop KMS	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-mapreduce	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-mapreduce-historyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	fil Hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-yarn-nodemanager	Suppléments	3.3.6. amzn.5-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Flink	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-ya rn-resour cemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-yarn-timelineserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hudi	Suppléments	0.15.0.am zn.0-1.am zn2023
Flink	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
Flink	Principal	hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
HBase	Principal	s3-dist-cp	Suppléments	2 h 33 - 1 h 00
HCatalog	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
HCatalog	Principal	mariadb-c onnector-java	Suppléments	
Hive	Principal	hadoop-httpfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Hive	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Hive	Principal	hive-hbase	Suppléments	3.1.3. amzn.12-1 . amzn2023

Application	Type de demande	Package	Type de package	Version du package
Hive	Principal	mariadb-connector-java	Suppléments	2.7.2-1
Hive	Principal	s3-dist-cp	Suppléments	2 h 33 - 1 h 00
Phoenix	Principal	s3-dist-cp	Suppléments	2 h 33 - 1 h 00
Spark	Principal	noyau de données Spark	Suppléments	3.5.1. amzn.0-1. amzn2023
Tez	Principal	hadoop-mapreduce	Suppléments	3.3.6. amzn.5-1. amzn2023
Tez	Principal	hadoop-mapreduce-historyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Tez	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	bigtop-utils	Suppléments	1.2.0-1. amzn2023
Trino	Principal	hadoop-client	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-hdfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-hdfs-datanode	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-hdfs-namenode	Suppléments	3.3.6. amzn.5-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Trino	Principal	Hadoop KMS	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-ma preduce	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-ma preduce-h istoryserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	fil Hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-yarn- nodemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-yarn- proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-ya rn-resour cemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-yarn- timelineserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hive	Suppléments	3.1.3. amzn.12-1 . amzn2023
Trino	Principal	hive-hcatalog- server	Suppléments	3.1.3. amzn.12-1 . amzn2023
Trino	Principal	mariadb-c onnector-java	Suppléments	2.7.2-1
Trino	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023

Application	Type de demande	Package	Type de package	Version du package
Trino	Principal	gardien de zoo	Suppléments	3.9.1. amzn.1-1. amzn2023
Trino	Principal	zookeeper-server	Suppléments	3.9.1. amzn.1-1. amzn2023
Trino	Principal	bigtop-groovy	Suppléments	2.5.4-1. amzn2023
Trino	Principal	bigtop-jsvc	Suppléments	1.2.4-1. amzn2023
Trino	Principal	hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-lzo	Suppléments	0.4.19-1 amzn2023
Trino	Principal	catalogue hive-h	Suppléments	3.1.3. amzn.12-1 . amzn2023
Trino	Principal	hive-jdbc	Suppléments	3.1.3. amzn.12-1 . amzn2023
Zookeeper	Principal	emrfs	Suppléments	2.64.0-1 amzn2023
Zookeeper	Principal	hadoop-client	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-hdfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-hdfs-datanode	Suppléments	3.3.6. amzn.5-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Zookeeper	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-hdfs-namenode	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-httpfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	Hadoop KMS	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-ma preduce	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-ma preduce-h istoryserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	fil Hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-yarn-nodemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-ya rn-resour cemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-yarn-timelineserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023

notes de mise à jour de la 7.4.0

Les notes de mise à jour suivantes incluent des informations relatives à la version 7.4.0 d'Amazon EMR.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 7.4.0 application upgrades include AWS SDK for Java 2.28.8, Delta 3.2.0-amzn-1, Flink 1.19.1, Hadoop 3.4.0, HBase 2.5.5, Hudi 0.15.0-amzn-1, Iceberg 1.6.1-amzn-0, Phoenix 5.2.0, Presto 0.287, Spark 3.5.2, Trino 446, and Zookeeper 3.9.2.
- Cette version ajoute 7 points de terminaison open source qui prennent en charge le chiffrement en transit sur le réseau. Cette amélioration inclut des frameworks open source tels que Hue, Livy JupyterEnterpriseGateway, Zookeeper et Ranger KMS. Si vous avez activé le chiffrement en transit et que vous souhaitez vous connecter à ces points de terminaison, vous devrez peut-être modifier votre code pour prendre en charge les communications TLS. Vous pouvez désactiver le protocole TLS pour n'importe lequel de ces points de terminaison selon vos besoins. Pour obtenir la liste complète des points de terminaison pris en charge, voir Comprendre le chiffrement en transit.

Problèmes connus

- La `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` propriété de Yarn ResourceManager, qui prenait auparavant en charge les gestionnaires de shuffle Hive depuis emr-6.11.0, applique désormais également une mise hors service progressive aux fichiers servis par le service de shuffle externe Spark, lors de l'exécution de Spark on Yarn dans emr-7.4.0.

  Notez que le comportement par défaut du service de shuffle externe Spark, lors de l'exécution de Spark on Yarn, est que le Yarn supprime les fichiers NodeManager de shuffle d'applications lorsque l'application est fermée. Cela peut avoir un impact sur la vitesse de mise hors service des nœuds et sur l'utilisation du calcul. Pour les applications de longue durée, pensez `spark.shuffle.service.removeShuffle true` à configurer pour supprimer les fichiers de shuffle qui ne sont plus utilisés, ce qui permet de mettre hors service plus rapidement les nœuds sans données de shuffle actives.
- Disponibilité d'Amazon S3 Select — Amazon S3 Select ne prend pas en charge les requêtes interrégionales pour Presto et Trino. Pour plus d'informations sur les fonctionnalités spécifiques disponibles pour les versions Presto et Trino, consultez Presto et Trino.
- Apache Ranger n'est pas pris en charge par la version 7.4 d'Amazon EMR.

- Apache Sqoop n'est pas compatible avec Apache Hadoop 3.4 en raison d'un problème connu, le SQOOP-3485. La version 7.4 d'Amazon EMR inclut mais ne prend pas en charge Apache Sqoop. Apache Sqoop sera supprimé des versions 7.5 et supérieures d'Amazon EMR.
- Lorsque le dimensionnement géré est activé, le YARN ResourceManager (RM) rencontre un problème de blocage critique, le rendant insensible lors de la transition de l'état DECOMMISSIONING à l'état DECOMMISSIONED avec des opérations simultanées. Cela se produit en particulier dans les environnements de grands clusters. Lorsque cela se produit, les transitions entre l'état du nœud et l'état de l'application ne sont pas traitées, paralysant ainsi les fonctionnalités de base du. ResourceManager Pour contourner le problème, redémarrez la JVM RM à l'aide de la commande suivante sur le nœud principal :

  ```
  sudo systemctl restart hadoop-yarn-resourcemanager
  ```

- Lorsqu'un cluster EMR est démarré avec des planificateurs YARN autres que le `CapacityScheduler` ou `FairScheduler` (par exemple, le planificateur FIFO), la `YARNMemoryAvailablePercentage` métrique n'est pas poussée vers. CloudWatch Ce problème a un impact sur la réduction d'échelle dans les clusters compatibles avec le dimensionnement géré. Cela peut également affecter le dimensionnement géré personnalisé s'il dépend de la `YARNMemoryAvailablePercentage` métrique. Pour contourner le problème, utilisez le `CapacityScheduler` ou`FairScheduler`.

Modifications, améliorations et problèmes résolus

- Amazon EMR utilise différentes versions de la TensorFlow bibliothèque en fonction des types d'instances que vous choisissez pour votre cluster. Amazon EMR prend également en charge les clusters dotés de types TensorFlow d'instances aarch64 (instances Graviton) à partir d'EMR-7.5.0 et versions ultérieures.
- Optimisation de la mise à l'échelle des demandes S3 pour les charges de travail Iceberg — Pour les clients qui utilisent Iceberg intégré aux versions d'EMR 34 v7.4, S3 a mis à jour le schéma d'entropie par défaut en utilisant un hachage de 20 caractères en base 2 (par exemple 35 01010110100110110010). Ce nouveau schéma d'entropie en base 2 réduit la plage de caractères et permet à S3 d'adapter automatiquement la capacité des demandes plus rapidement aux exigences de la charge de travail cible d'Iceberg. Cela réduit également le temps pendant lequel une charge de travail observe 503 réponses d'accélérateur.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250527.1	6.1.134-152.225 mai 2023	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique (Taipei), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250512,(	6.1.134-152.225 mai 2023	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Bahreïn) (Émirats arabes unis), Canada (centre), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			, AWS GovCloud (États-Unis est), AWS GovCloud (États-Unis ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250428.1	6,1134-150,224 mai 2023	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Malaisie), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250414,(	6.1.132-147.221 mai 2023	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7.2 250331.(	6.1.131-143.221 mai 2023	18 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre), Asie-Pacifique (Tokyo)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250303.(	6.1.129-138.220 août 2023	27 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250218,:	6.1.128-136.201. mai 2023	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250211.(	6.1.127-135.201. mai 2023	26 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250123.4	6.1.124-134.200. mai 2023	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250115,(	6.1.119-129.201.am zn2023	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 241121.(	6.1.115-126.197. mai 2023	12 décembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5.2 240916	6 janvier 109 à 118,189 mai 2023	6 septembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

Versions Java 7.4.0 par défaut

Les versions 7.4 et supérieures d'Amazon EMR sont fournies par défaut avec Amazon Corretto 17 (basé sur OpenJDK) pour les applications compatibles avec Corretto 17 (JDK 17), à l'exception d'Apache Livy.

Le tableau suivant indique les versions Java par défaut pour les applications dans Amazon EMR 7.4.0. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	21, 17 ANS
Zeppelin	17, 11, 8
Pig	8
Zookeeper	8

Versions des composants 7.4.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
adot-java-agent	1,31,0	Agent Java qui collecte des métriques à partir des démons d'application.
delta	3.2.0-amzn-1	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-amazon-cloudwatch-agent	1,300032,2-amzn-0	Une application qui collecte des métriques internes au niveau du système et des métriques d'applica tion personnalisées à partir d'instances Amazon EC2 .
emr-ddb	5.6.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3,13,0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3,17,0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.17.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.34.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.13.0	Connecteur S3Select EMR
emr-wal-cli	1.3.1	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,65,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.19.1-amzn-0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.19.1-amzn-0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
hadoop-client	3.4.0-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.4.0-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.4.0-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.4.0-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	3.4.0-amzn-0	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.4.0-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.4.0-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.4.0-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.4.0-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.4.0-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.4.0-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2,5,5-amzn-0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	2,5,5-amzn-0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2,5,5-amzn-0	HBase client en ligne de commande.
hbase-rest-server	2,5,5-amzn-0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2,5,5-amzn-0	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2,5,5-amzn-0	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-13	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-13	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-13	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-13	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-13	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	3.1.3-amzn-13	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-13	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0,15,0-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0,15,0-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0,15,0-amzn-1	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0,15,0-amzn-1	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1.6.1-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.

Composant	Version	Description
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.8.0 - incubation	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	12.3.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.2.0	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	6.0.0	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica

Composant	Version	Description
presto-coordinator	0,287-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0,287-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0,287-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	446-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	446-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	446-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.3.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger

Composant	Version	Description
spark-client	3.5.2-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.5.2-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.5.2-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.5.2-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	24.08.1-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.16.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-11	Les bibliothèques et l'applica tion tez YARN.
tez-on-worker	0.10.2-amzn-11	Les bibliothèques et l'applica tion tez YARN pour les nœuds de travail.
zeppelin-server	0.11.1	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.9.2-amzn-0	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.9.2-amzn-0	ZooKeeper client en ligne de commande.

7.4.0 classifications de configuration

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

classifications EMR-7.4.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classifications	Description	Actions de reconfiguration
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.

Classifications	Description	Actions de reconfiguration
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.pr operties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-s ite.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.pr operties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.prope rties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_c onfig.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistan ce S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.

Classifications	Description	Actions de reconfiguration
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
phoenix-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.p roperties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.

Classifications	Description	Actions de reconfiguration
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Modifiez les valeurs dans ZooKeeper le fichier logback.xml.	Restarts Zookeeper server.
emr-metrics	Modifiez les paramètres métriques EMR pour ce nœud.	Restarts the CloudWatchAgent service.

journal des modifications 7.4.0

Journal des modifications pour la version 7.4.0 et notes de version

Date	Événement	Description
18/11/2024	Publication de documents	Notes de mise à jour d'Amazon EMR 7.4.0 publiées pour la première fois
13/11/2024	Déploiement terminé	Amazon EMR 7.4.0 entièrement déployé dans toutes les régions prises en charge
31/10/2024	Début du déploiement	Amazon EMR 7.4.0 a été déployé pour la première fois dans les régions commerciales initiales

Amazon EMR version 7.3.0

Cycle de vie pris en charge en 7.3.0

Le tableau suivant décrit les dates de cycle de vie prises en charge pour la version 7.3.0 d'Amazon EMR.

Phase de support	Date	
Date de sortie initiale	16 octobre 2024	
Support standard jusqu'à	15 octobre 2026	
Fin de l'assistance	16 octobre 2026	
Fin de vie	16 octobre 2027	

Pour obtenir une liste des vulnérabilités et des risques courants critiques et élevés (CVEs) qui n'affectent pas les clusters EMR dans le cadre des logiciels et des configurations recommandés, voir 7.3.0 Vulnérabilités et expositions courantes connues pour les moteurs principaux.

composants de la version 7.3.0

Amazon EMR 7.3.0 inclut les applications suivantes :

Moteurs principaux :

AmazonCloudWatchAgent, Delta. Flink, Hadoop, Hive HBase, Hudi HCatalog, Livy, Iceberg, Phoenix, Spark, Tez, Trino JupyterEnterpriseGateway, gardien de zoo

Suppléments :

Hue JupyterHub, Oozie, Pig, Presto, Sqoop, Zeppelin TensorFlow

Pour plus d'informations sur les packages inclus, consultez la section Versions des applications dans les versions 7.x d'Amazon EMR. Pour plus d'informations sur les moteurs principaux et les options supplémentaires, consultez Amazon EMR Standard Support.

Versions de l'application 7.3.0

Cette version inclut les applications suivantes : AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, Oozie, PhoenixPig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS SDK pour Java	2,25,70, 1,12,747	2,23,18, 1,12,705	2,23,18, 1,12,656	2.20.160-amzn-0, 1.12.569
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3.9
Scala	2,1,18	2,1,17	2,1,17	2,1,17
AmazonCloudWatchAgent	1,300032,2	1,300032,2	1,300032,2	1,300031,1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18.1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4,17	2.4,17	2.4,17	2.4,17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0,15,0-amzn-0	0,14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1,5.0-amzn-0	1.4.3-amzn-0	1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,285	0,285	0,284	0,283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

7.3.0 vulnérabilités et expositions courantes

Le tableau suivant répertorie tout CVEs ce qui n'a pas d'impact sur les clusters EMR qui s'exécutent sur les configurations recommandées d'Amazon EMR 7.3.0. Amazon EMR dépend de l'open source en amont pour la disponibilité des correctifs et fournira la dernière version stable dans le cadre des composants du moteur principal dans les 90 jours suivant la vérification des correctifs par Amazon EMR.

Ce tableau a été mis à jour le 4 octobre 2024.

ID CVE	Sévérité	URL des détails du CVE
CVE-2018-1000180	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-1000180
CVE-2020-28052	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-28052
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2021-31684	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2_1471	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_1471
CVE-2_25857	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25857
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2_40149	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40149
CVE-2_40150	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40150
CVE-2_40152	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40152

ID CVE	Sévérité	URL des détails du CVE
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-avril 2004
CVE-2_45685	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45685
CVE-2_45693	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45693
CVE-2023-1370	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-1436	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2023-39410	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-44981	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2023-50387	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-50387

ID CVE	Sévérité	URL des détails du CVE
CVE-2012-0881	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2012-0881
CVE-2013-4002	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2015-1832	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2015-5237	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2018-11307	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14719
CVE-2018-14720	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19361

ID CVE	Sévérité	URL des détails du CVE
CVE-2018-19362	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19362
CVE-2018-7489	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-7489
CVE-2019-12086	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2019-14379	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2019-17267	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17531

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-20330	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-13936	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2020-13949	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13949

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-14060	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-25649	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2020-35490	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36181

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-36182	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9548

ID CVE	Sévérité	URL des détails du CVE
CVE-2021-20190	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2021-31684	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2021-34538	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-34538	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-35515	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2_25647	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25647
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2_36364	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_36364

ID CVE	Sévérité	URL des détails du CVE
CVE-2_37865	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_37865
CVE-2_37866	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_37866
CVE-2_39135	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_39135
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004
CVE-2_46337	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_46337
CVE-2_46751	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_46751
CVE-2023-1370	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34454

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-34455	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-39410	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2016-5018	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2016-6796	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-02-05
CVE-2020-13949	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-43642	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2015-2156	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2015-2156
CVE-2017-18640	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-18640
CVE-2017-18640	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-18640
CVE-2019-16869	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-16869

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-20444	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20444
CVE-2019-20445	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20445
CVE-2021-37136	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2021-37137	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2_1471	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_1471
CVE-2_1471	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_1471
CVE-2_25857	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25857
CVE-2_25857	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25857
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2_40149	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40149
CVE-2_40150	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40150

ID CVE	Sévérité	URL des détails du CVE
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004
CVE-2_45685	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45685
CVE-2_45693	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45693
CVE-2023-1436	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2023-44981	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2023-44981	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2019-0205	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-02-05
CVE-2020-13949	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2021-34538	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-34538

ID CVE	Sévérité	URL des détails du CVE
CVE-2_45688	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45688
CVE-2_46337	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_46337
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-3635	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-5072	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2024-21634	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2_40149	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40149
CVE-2_40150	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40150
CVE-2_45685	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45685
CVE-2_45693	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45693
CVE-2023-1436	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-6378	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6378	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6378

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-6481	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2023-6481	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6481

Détails des composants de la version 7.3.0

Consultez le tableau suivant pour plus d'informations sur les packages Extras dans Amazon EMR 7.3.0.

Application	Type de demande	Package	Type de package	Version du package
AmazonCloudWatchAgent	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.5-1. amzn2023
AmazonCloudWatchAgent	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
AmazonCloudWatchAgent	Principal	s3-dist-cp	Suppléments	2 h 33 - 1 h 00
AmazonCloudWatchAgent	Principal	gardien de zoo	Suppléments	3.9.1. amzn.1-1. amzn2023
AmazonCloudWatchAgent	Principal	zookeeper-server	Suppléments	3.9.1. amzn.1-1. amzn2023
Flink	Principal	hadoop-client	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-hdfs	Suppléments	3.3.6. amzn.5-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Flink	Principal	hadoop-hdfs-datanode	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-hdfs-namenode	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-httpfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	Hadoop KMS	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-ma preduce	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-ma preduce-h istoryserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	fil Hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-yarn-nodemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hadoop-ya rn-resour cemanager	Suppléments	3.3.6. amzn.5-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Flink	Principal	hadoop-yarn-timelineserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Flink	Principal	hudi	Suppléments	0.15.0.am zn.0-1.am zn2023
Flink	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
Flink	Principal	hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
HBase	Principal	s3-dist-cp	Suppléments	2 h 33 - 1 h 00
HCatalog	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
HCatalog	Principal	mariadb-c onnector-java	Suppléments	
Hive	Principal	hadoop-httpfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Hive	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Hive	Principal	hive-hbase	Suppléments	3.1.3. amzn.12-1 . amzn2023
Hive	Principal	mariadb-c onnector-java	Suppléments	2.7.2-1
Hive	Principal	s3-dist-cp	Suppléments	2 h 33 - 1 h 00
Phoenix	Principal	s3-dist-cp	Suppléments	2 h 33 - 1 h 00

Application	Type de demande	Package	Type de package	Version du package
Spark	Principal	noyau de données Spark	Suppléments	3.5.1. amzn.0-1. amzn2023
Tez	Principal	hadoop-ma preduce	Suppléments	3.3.6. amzn.5-1. amzn2023
Tez	Principal	hadoop-ma preduce-h istoryserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Tez	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	bigtop-utils	Suppléments	1.2.0-1. amzn2023
Trino	Principal	hadoop-client	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-hdfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-hdfs-datanode	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-hdfs-namenode	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	Hadoop KMS	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-ma preduce	Suppléments	3.3.6. amzn.5-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Trino	Principal	hadoop-mapreduce-historyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	fil Hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-yarn-nodemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-yarn-resourcemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-yarn-timelineserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hive	Suppléments	3.1.3. amzn.12-1. amzn2023
Trino	Principal	hive-hcatalog-server	Suppléments	3.1.3. amzn.12-1. amzn2023
Trino	Principal	mariadb-connector-java	Suppléments	2.7.2-1
Trino	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
Trino	Principal	gardien de zoo	Suppléments	3.9.1. amzn.1-1. amzn2023
Trino	Principal	zookeeper-server	Suppléments	3.9.1. amzn.1-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Trino	Principal	bigtop-groovy	Suppléments	2.5.4-1. amzn2023
Trino	Principal	bigtop-jsvc	Suppléments	1.2.4-1. amzn2023
Trino	Principal	hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
Trino	Principal	hadoop-lzo	Suppléments	0.4.19-1 amzn2023
Trino	Principal	catalogue hive-h	Suppléments	3.1.3. amzn.12-1 . amzn2023
Trino	Principal	hive-jdbc	Suppléments	3.1.3. amzn.12-1 . amzn2023
Zookeeper	Principal	emrfs	Suppléments	2.64.0-1 amzn2023
Zookeeper	Principal	hadoop-client	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-hdfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-hdfs-datanode	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-hdfs-namenode	Suppléments	3.3.6. amzn.5-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Zookeeper	Principal	hadoop-httpfs	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	Hadoop KMS	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-ma preduce	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-ma preduce-h istoryserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	fil Hadoop	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-yarn- nodemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-yarn- proxyserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-ya rn-resour cemanager	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	hadoop-yarn- timelineserver	Suppléments	3.3.6. amzn.5-1. amzn2023
Zookeeper	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023

notes de mise à jour de la 7.3.0

Les notes de mise à jour suivantes incluent des informations relatives à la version 7.3.0 d'Amazon EMR. Les modifications concernent la version 7.3.0.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 7.3.0 application upgrades include Delta 3.2.0-amzn-0, Hudi 0.15.0-amzn-0, TensorFlow 2.16.1, Trino 442, and Zeppelin 0.11.1.
- Cette version ajoute 22 points de terminaison open source qui prennent en charge le chiffrement en transit sur le réseau. Cette amélioration inclut des frameworks open source tels que Hadoop, Hive, Tez HBase, Phoenix et Flink. Si vous avez activé le chiffrement en transit et que vous souhaitez vous connecter à ces points de terminaison, vous devrez peut-être modifier votre code pour prendre en charge les communications TLS. Vous pouvez désactiver le protocole TLS pour n'importe lequel de ces points de terminaison selon vos besoins. Pour obtenir la liste complète des points de terminaison pris en charge, voir Comprendre le chiffrement en transit.
- Avec Amazon EMR 7.3.0, Apache Hive with Iceberg prend désormais en charge le catalogue de données AWS Glue en tant que métastore, ce qui vous offre des avantages tels qu'un référentiel centralisé pour stocker les métadonnées, explorer automatiquement les sources de données et extraire les métadonnées, et l'intégration avec d'autres services pour des flux de travail fluides.

Problèmes connus

- Python 3.11 n'est pas pris en charge avec EMR Studio.
- Si vous utilisez le chiffrement en transit et l'authentification Kerberos, vous ne pourrez pas utiliser les applications Spark qui dépendent du métastore Hive. Ce problème est résolu dans HIVE-16340 et sera entièrement résolu dans SPARK-44114 lorsque Spark pourra passer à Hive 3 en open source. En attendant, vous pouvez contourner temporairement ce problème en réglant `hive.metastore.use.SSL` sur`false`. Pour plus d'informations, consultez Configuration des applications.
  - Si vous n'utilisez pas l'authentification Kerberos, vous devez configurer Spark pour qu'il utilise le protocole SSL. Pour ce faire, définissez les configurations suivantes lorsque vous provisionnez votre cluster.

```
[{
    "Classification": "spark-defaults",
    "Properties": {
        "spark.hadoop.hive.metastore.truststore.path": "/usr/share/aws/emr/
security/conf/keystore-thrift.jks",
        "spark.hadoop.hive.metastore.use.SSL": "true"
    }
}]
```

- Si vous utilisez des domaines personnalisés, certains liens de l'interface utilisateur de Spark History Server dans la AWS console risquent de ne pas fonctionner. Téléchargez le script suivant dans un compartiment S3, puis ajoutez une étape pour utiliser le script shell.

```
#!/bin/bash

if grep -q ":8042" /etc/nginx/nginx.conf; then
    sudo sed -i '=s#\^(\\/nm\\/(?<nodeAddress>.*:8042)\\/(?<url>.*)\$)#^(\\/nm\\/
(?<nodeAddress>(\\.|-|[a-zA-Z0-9])*:8042)\\/(?<url>(node|conf|logs|stacks|jmx).*)\
$)#g' /etc/nginx/nginx.conf
else
    sudo sed -i 's#\^(\\/nm\\/(?<nodeAddress>.*:8044)\\/(?<url>.*)\$)#^(\\/nm\\/
(?<nodeAddress>(\\.|-|[a-zA-Z0-9])*:8044)\\/(?<url>(node|conf|logs|stacks|jmx).*)\
$)#g' /etc/nginx/nginx.conf
fi

sudo systemctl restart nginx
```

Modifications, améliorations et problèmes résolus

- Pour prendre en charge la vérification du nom d'hôte TLS pour YARN, cette version modifie la `yarn.resourcemanager.hostname` configuration de l'adresse IP de l'instance principale au nom DNS privé de l'instance principale.
- MXNet le projet a atteint la fin de sa durée de vie et est maintenant archivé. the section called “emr-7.2.0”est la dernière version qui inclut MXNet.

- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250527.1	6.1.134-152.225 mai 2023	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250512,(	6.1.134-152.225 mai 2023	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Bahreïn) (Émirats arabes unis), Canada (centre), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			, AWS GovCloud (États-Unis est), AWS GovCloud (États-Unis ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250428.1	6,1134-150,224 mai 2023	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Malaisie), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250414,(	6.1.132-147.221 mai 2023	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7.2 250331.(	6.1.131-143.221 mai 2023	18 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre), Asie-Pacifique (Tokyo)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250303.(	6.1.129-138.220 août 2023	27 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250218,:	6.1.128-136.201. mai 2023	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250211.(	6.1.127-135.201. mai 2023	26 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250123.4	6.1.124-134.200. mai 2023	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250115,(	6.1.119-129.201.am zn2023	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 241121.(	6.1.115-126.197. mai 2023	12 décembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 241031.(	6.1.112-124.190. mai 2023	15 novembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5.2 240916	6 janvier 109 à 118,189 mai 2023	6 septembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

Versions Java 7.3.0 par défaut

Les versions 7.3 et supérieures d'Amazon EMR sont fournies par défaut avec Amazon Corretto 17 (basé sur OpenJDK) pour les applications compatibles avec Corretto 17 (JDK 17), à l'exception d'Apache Livy.

Le tableau suivant indique les versions Java par défaut pour les applications dans Amazon EMR 7.3.0. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	21, 17 ANS
Zeppelin	8
Pig	8
Zookeeper	17, 11, 8

Versions des composants 7.3.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
adot-java-agent	1,31,0	Agent Java qui collecte des métriques à partir des démons d'application.
delta	3.2.0-amzn-0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-amazon-cloudwatch-agent	1,300032,2-amzn-0	Une application qui collecte des métriques internes au niveau du système et des métriques d'applica tion personnalisées à partir d'instances Amazon EC2 .
emr-ddb	5.5.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.12.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3,16,0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.16.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.33.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.12.0	Connecteur S3Select EMR
emr-wal-cli	1.3.0	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,64,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1,18.1-amzn-2	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1,18.1-amzn-2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
hadoop-client	3.3.6-amzn-5	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.3.6-amzn-5	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.6-amzn-5	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.3.6-amzn-5	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	3.3.6-amzn-5	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.6-amzn-5	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.6-amzn-5	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-5	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.6-amzn-5	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.3.6-amzn-5	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.3.6-amzn-5	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2,4,17-amzn-7	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	2,4,17-amzn-7	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2,4,17-amzn-7	HBase client en ligne de commande.
hbase-rest-server	2,4,17-amzn-7	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2,4,17-amzn-7	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2,4,17-amzn-7	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-12	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-12	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-12	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-12	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-12	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	3.1.3-amzn-12	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-12	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0,15,0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0,15,0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0,15,0-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0,15,0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1.5.2-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.

Composant	Version	Description
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.8.0 - incubation	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	12.3.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.3	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	6.0.0	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica

Composant	Version	Description
presto-coordinator	0,285-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0,285-amzn-1	Service pour exécuter les éléments d'une requête.
presto-client	0,285-amzn-1	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	442-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	442-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	442-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.3.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger

Composant	Version	Description
spark-client	3.5.1-amzn-1	Clients de ligne de commande Spark.
spark-history-server	3.5.1-amzn-1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.5.1-amzn-1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.5.1-amzn-1	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	24.06.1-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.16.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-10	Les bibliothèques et l'applica tion tez YARN.
tez-on-worker	0.10.2-amzn-10	Les bibliothèques et l'applica tion tez YARN pour les nœuds de travail.
zeppelin-server	0.11.1	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.9.1-amzn-1	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.9.1-amzn-1	ZooKeeper client en ligne de commande.

7.3.0 classifications de configuration

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

classifications EMR-7.3.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classifications	Description	Actions de reconfiguration
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.

Classifications	Description	Actions de reconfiguration
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.pr operties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-s ite.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.pr operties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.prope rties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_c onfig.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistan ce S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.

Classifications	Description	Actions de reconfiguration
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.p roperties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.

Classifications	Description	Actions de reconfiguration
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Modifiez les valeurs dans ZooKeeper le fichier logback.xml.	Restarts Zookeeper server.
emr-metrics	Modifiez les paramètres métriques EMR pour ce nœud.	Restarts the CloudWatchAgent service.

journal des modifications 7.3.0

Journal des modifications pour la version 7.3.0 et notes de version

Date	Événement	Description
17/10/2024	Publication de documents	Notes de mise à jour d'Amazon EMR 7.3.0 publiées pour la première fois
16/10/2024	Déploiement terminé	Amazon EMR 7.3.0 entièrement déployé dans toutes les régions prises en charge
14-10-2024	Première version	Amazon EMR 7.3.0 a été déployé pour la première fois dans les régions commerciales initiales

Amazon EMR version 7.2.0

Cycle de vie pris en charge en 7.2.0

Le tableau suivant décrit les dates de cycle de vie prises en charge pour la version 7.2.0 d'Amazon EMR.

Phase de support	Date
Date de sortie initiale	25 juillet 2024
Support standard jusqu'à	24 juillet 2026
Fin de l'assistance	25 juillet 2026
Fin de vie	25 juillet 2027

Pour obtenir une liste des vulnérabilités et des risques courants critiques et élevés (CVEs) qui n'affectent pas les clusters EMR dans le cadre des logiciels et des configurations recommandés, voir

7.2.0 Vulnérabilités et expositions courantes connues pour les moteurs principaux. Pour une liste de ceux CVEs qui ont été corrigés dans la version 7.2.0, voir 7.2.0 a corrigé les vulnérabilités et les expositions courantes.

composants de la version 7.2.0

Amazon EMR 7.2.0 inclut les applications suivantes :

Moteurs principaux :

AmazonCloudWatchAgent, Delta. Flink, Hadoop, Hive HBase, Hudi HCatalog, Livy, Iceberg, Phoenix, Spark, Tez, Trino JupyterEnterpriseGateway, gardien de zoo

Suppléments :

Hue JupyterHub, Oozie MXNet, Pig, Presto, Sqoop, Zeppelin TensorFlow

Pour plus d'informations sur les packages inclus, consultez la section Versions des applications dans les versions 7.x d'Amazon EMR. Pour plus d'informations sur les moteurs principaux et les options supplémentaires, consultez Amazon EMR Standard Support.

Versions de l'application 7.2.0

Cette version inclut les applications suivantes : AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, OoziePhoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS SDK pour Java	2,25,70, 1,12,747	2,23,18, 1,12,705	2,23,18, 1,12,656	2.20.160-amzn-0, 1.12.569
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3.9
Scala	2,1,18	2,1,17	2,1,17	2,1,17
AmazonCloudWatchAgent	1,300032,2	1,300032,2	1,300032,2	1,300031,1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18.1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4,17	2.4,17	2.4,17	2.4,17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0,15,0-amzn-0	0,14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1,5.0-amzn-0	1.4.3-amzn-0	1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,285	0,285	0,284	0,283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

7.2.0 vulnérabilités et risques courants

Cette page répertorie toutes les vulnérabilités et expositions courantes (CVEs) liées à Amazon EMR 7.2.0.

7.2.0 vulnérabilités et risques courants connus pour les moteurs principaux

Le tableau suivant répertorie tout CVEs ce qui n'a aucun impact sur les clusters EMR qui s'exécutent sur les configurations recommandées d'Amazon EMR 7.2.0. Amazon EMR dépend de l'open source en amont pour la disponibilité des correctifs et fournira la dernière version stable dans le cadre des composants du moteur principal dans les 90 jours suivant la vérification des correctifs par Amazon EMR.

Ce tableau a été mis à jour le 16 septembre 2024.

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2021-31684	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2_1471	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_1471
CVE-2_25857	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25857
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2_40149	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40149
CVE-2_40150	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40150
CVE-2_40152	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40152
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003

ID CVE	Sévérité	URL des détails du CVE
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004
CVE-2_45685	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45685
CVE-2_45693	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45693
CVE-2023-1370	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-1436	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2023-39410	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-44981	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2015-5237	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2017-15095	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2017-17485	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-17485

ID CVE	Sévérité	URL des détails du CVE
CVE-2017-7525	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-7525
CVE-2018-11307	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14719
CVE-2018-14720	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19362
CVE-2018-5968	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2018-7489	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-7489

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-10172	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-10172
CVE-2019-10202	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-10202
CVE-2019-12086	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2019-14379	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2019-17267	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17531

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-20330	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-14060	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14061

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-14062	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-35490	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36184

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-36185	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2021-31684	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2021-33813	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-33813

ID CVE	Sévérité	URL des détails du CVE
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004
CVE-2023-1370	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-25194	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34610	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34610

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-3635	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-39410	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-44981	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2012-0881	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2012-0881
CVE-2013-4002	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2015-1832	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2015-5237	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2018-11307	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14719

ID CVE	Sévérité	URL des détails du CVE
CVE-2018-14720	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19362
CVE-2018-7489	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-7489
CVE-2019-12086	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2019-14379	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16335

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-16942	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2019-17267	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17531
CVE-2019-20330	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11113

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-11619	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-13936	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2020-13949	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-14060	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-25649	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2020-35490	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35491

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-35728	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36518

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-8840	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2021-31684	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2021-34538	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-34538	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-35515	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2_25647	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25647

ID CVE	Sévérité	URL des détails du CVE
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2_36364	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_36364
CVE-2_37865	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_37865
CVE-2_37866	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_37866
CVE-2_39135	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_39135
CVE-2_40149	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40149
CVE-2_40150	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40150
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004

ID CVE	Sévérité	URL des détails du CVE
CVE-2_45685	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45685
CVE-2_45693	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45693
CVE-2_46337	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_46337
CVE-2_46751	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_46751
CVE-2023-1370	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-1436	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-39410	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-39410

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-43642	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2_25647	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25647
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-2_42889	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_42889
CVE-2_46751	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_46751
CVE-2023-1370	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-20883	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-20883
CVE-2023-25194	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34455

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-39410	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-39410	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-39410	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-5072	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2018-8768	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-8768
CVE-2021-32797	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-32797
CVE-2021-32798	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-32798
CVE-2_24758	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_24758
CVE-2023-24580	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-24580
CVE-2023-36053	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-36053
CVE-2023-41164	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-41164
CVE-2023-43665	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-43665

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-46695	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-46695
CVE-2024-24680	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-24680
CVE-2016-5018	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2016-6796	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-02-05
CVE-2020-13949	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-43642	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2012-0881	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2012-0881
CVE-2013-4002	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2019-17566	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-17566
CVE-2019-17566	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-17566
CVE-2020-11988	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11988

ID CVE	Sévérité	URL des détails du CVE
CVE-2021-33813	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-33813
CVE-2_25647	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25647
CVE-2_34169	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_34169
CVE-2_40146	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40146
CVE-2_41704	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_41704
CVE-2_42890	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_42890
CVE-2_44729	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_44729
CVE-2_44729	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_44729
CVE-2_44729	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_44729
CVE-2_46337	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_46337
CVE-2023-4759	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-4759
CVE-2015-2156	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2015-2156
CVE-2017-18640	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-18640

ID CVE	Sévérité	URL des détails du CVE
CVE-2017-18640	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-18640
CVE-2019-16869	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-16869
CVE-2019-20444	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20444
CVE-2019-20445	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20445
CVE-2021-37136	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2021-37137	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2_1471	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_1471
CVE-2_1471	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_1471
CVE-2_25857	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25857
CVE-2_25857	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25857
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510

ID CVE	Sévérité	URL des détails du CVE
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2023-44981	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2023-44981	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2015-2080	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2015-2080
CVE-2015-4852	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2015-4852
CVE-2015-6420	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2015-6420
CVE-2015-7501	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2015-7501
CVE-2017-1000034	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-1000034
CVE-2017-12612	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-12612

ID CVE	Sévérité	URL des détails du CVE
CVE-2017-15288	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-15288
CVE-2017-5637	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-5637
CVE-2017-7656	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-7656
CVE-2017-7657	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-7657
CVE-2017-7658	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-7658
CVE-2017-7687	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-7687
CVE-2017-9735	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-9735
CVE-2017-9790	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2017-9790
CVE-2018-11793	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-11793
CVE-2018-11804	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-11804
CVE-2018-17190	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-17190
CVE-2018-8012	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-8012
CVE-2019-0204	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-02-04

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-10099	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-10099
CVE-2020-17533	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-17533
CVE-2020-9480	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9480
CVE-2021-28165	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-28165
CVE-2021-35515	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35515	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35516	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-35517	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2021-36090	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2_25647	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25647

ID CVE	Sévérité	URL des détails du CVE
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2_37866	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_37866
CVE-2_46751	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_46751
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-43642	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-44981	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2014-0114	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2014-01-14
CVE-2016-1000338	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-1000338

ID CVE	Sévérité	URL des détails du CVE
CVE-2016-1000340	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-1000340
CVE-2016-1000342	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-1000342
CVE-2016-1000343	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-1000343
CVE-2016-1000344	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-1000344
CVE-2016-1000352	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2016-1000352
CVE-2017-1000487	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-1000487
CVE-2017-12629	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-12629
CVE-2018-10936	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-10936
CVE-2018-3258	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-3258
CVE-2018-3827	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-3827
CVE-2018-3831	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-3831
CVE-2019-0205	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-02-05
CVE-2019-10086	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-10086

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-10086	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-10086
CVE-2019-7611	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-7611
CVE-2020-13692	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13692
CVE-2020-13949	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-28491	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-28491
CVE-2020-7692	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-7692
CVE-2021-22573	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-22573
CVE-2021-22573	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-22573
CVE-2021-25646	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-25646
CVE-2021-26291	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-26291
CVE-2021-26291	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-26291
CVE-2021-26919	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-26919
CVE-2021-35515	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35515

ID CVE	Sévérité	URL des détails du CVE
CVE-2021-35516	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2_1471	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_1471
CVE-2_25647	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25647
CVE-2_25857	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25857
CVE-2_31197	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_31197
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_33980	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_33980
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510

ID CVE	Sévérité	URL des détails du CVE
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004
CVE-2_4244	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_4244
CVE-2_45688	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45688
CVE-2023-1370	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-1428	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1428
CVE-2023-25194	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-31418	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-31418
CVE-2023-32731	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-32731

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-34453	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-3635	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-38647	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-38647
CVE-2023-39410	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-44981	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2023-46120	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-46120
CVE-2023-5072	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2023-6378	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2024-1597	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-1597

ID CVE	Sévérité	URL des détails du CVE
CVE-2024-1597	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-1597
CVE-2021-35958	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-35958
CVE-2019-0205	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-02-05
CVE-2020-13949	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2021-34538	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2_45688	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45688
CVE-2_46337	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_46337
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-3635	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-5072	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2024-21634	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2021-35515	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35516

ID CVE	Sévérité	URL des détails du CVE
CVE-2021-35517	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2_40149	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40149
CVE-2_40150	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40150
CVE-2_45685	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45685
CVE-2_45693	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_45693
CVE-2023-1436	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2014-0114	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2014-01-14
CVE-2015-2156	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2015-2156
CVE-2018-10054	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-10054
CVE-2019-0205	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-02-05
CVE-2019-10086	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-10086
CVE-2019-10086	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-10086

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-16869	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-16869
CVE-2019-17195	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17195
CVE-2019-20444	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20444
CVE-2019-20445	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20445
CVE-2020-13949	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-9492	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-9492
CVE-2021-37136	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2021-37137	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2021-37404	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2021-37404
CVE-2_25168	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_25168
CVE-2_25647	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25647
CVE-2_25647	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_25647
CVE-2_26612	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2_26612

ID CVE	Sévérité	URL des détails du CVE
CVE-2_3171	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3171
CVE-2_3509	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3509
CVE-2_3510	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_3510
CVE-2_40152	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2_40152
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-3635	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-3635	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-38647	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2023-38647
CVE-2023-43642	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-43642

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-6378	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2024-1597	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2024-1597
CVE-2017-15095	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2017-17485	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2018-11307	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14719
CVE-2018-14720	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19360

ID CVE	Sévérité	URL des détails du CVE
CVE-2018-19361	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-19362
CVE-2018-5968	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2018-7489	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2018-7489
CVE-2019-12086	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2019-12086	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2019-14379	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2019-14379	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2019-14439	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2019-14540	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14892

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-14892	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2019-16335	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2019-16942	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2019-16943	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2019-17195	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17195
CVE-2019-17267	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2019-17267	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17531
CVE-2019-17531	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-17531

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-20330	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2019-20330	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10650	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-10672	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-10673	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-10968	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2020-10969	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11111

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-11111	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-11620	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-14060	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14060	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14061	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14062	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14195

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-14195	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-24616	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-24750	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-25649	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2020-25649	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2020-25649	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2020-35211	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35211
CVE-2020-35490	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35490	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35490	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35491

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-35491	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35491	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-35728	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-35728	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36179	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36179	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36180	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36180	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36181	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36181

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-36181	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36182	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36182	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36183	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36183	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36184	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36184	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36185	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36185	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36185

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-36186	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36186	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36186	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36187	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36187	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36188	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36188	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36189	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36189	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36518

ID CVE	Sévérité	URL des détails du CVE
CVE-2020-36518	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-36518	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-36518	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-7692	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-7692
CVE-2020-8840	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-8840	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9546	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9547	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2020-9548	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-20190

ID CVE	Sévérité	URL des détails du CVE
CVE-2021-20190	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2021-20190	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2021-22573	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-22573
CVE-2021-33036	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-33036
CVE-2021-35515	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35515	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35516	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-37714	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37714
CVE-2021-37714	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37714
CVE-2021-37714	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37714
CVE-2021-37714	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-37714
CVE-2021-46877	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-46877

ID CVE	Sévérité	URL des détails du CVE
CVE-2_25647	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-02:25647
CVE-2_32532	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-432 32532
CVE-2_40149	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-3_40149
CVE-2_40150	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-432 40150
CVE-2_40664	CRITIQUE	https://nvd.nist. gov/vuln/detail/CVE-40664
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-2_42003	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 avril 2003
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004
CVE-04-2004	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-4 février 2004

ID CVE	Sévérité	URL des détails du CVE
CVE-2_45685	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-3_45685
CVE-2_45688	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-45688
CVE-2_45693	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-3_45693
CVE-2023-1436	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2_40149	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-3_40149
CVE-2_40150	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-432 40150
CVE-2_45685	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-432 45685
CVE-2_45693	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-45693
CVE-2023-1436	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-6378	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6378	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6378

ID CVE	Sévérité	URL des détails du CVE
CVE-2023-6481	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2023-6481	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2023-6481

7.2.0 a corrigé des vulnérabilités et des risques courants pour les moteurs principaux

Le tableau suivant répertorie toutes les corrections apportées CVEs dans Amazon EMR 7.2.0.

ID CVE	Sévérité	URL des détails du CVE
CVE-2019-12402	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2019-12402
CVE-2021-41561	ÉLEVÉ	https://nvd.nist. gov/vuln/detail/CVE-2021-41561

Détails des composants de la version 7.2.0

Consultez le tableau suivant pour plus d'informations sur les packages Extras dans Amazon EMR 7.2.0.

Application	Type de demande	Package	Type de package	Version du package
AmazonCloudWatchAgent	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.4-1. amzn2023
AmazonCloudWatchAgent	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
AmazonCloudWatchAgent	Principal	s3-dist-cp	Suppléments	2.32.0-1 amzn2023

Application	Type de demande	Package	Type de package	Version du package
AmazonCloudWatchAgent	Principal	gardien de zoo	Suppléments	3.9.1. amzn.1-1. amzn2023
AmazonCloudWatchAgent	Principal	zookeeper-server	Suppléments	3.9.1. amzn.1-1. amzn2023
Flink	Principal	hadoop-client	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	hadoop-hdfs	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	hadoop-hdfs-datanode	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	hadoop-hdfs-namenode	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	hadoop-httpfs	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	Hadoop KMS	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	hadoop-mapreduce	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	hadoop-mapreduce-historyserver	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	fil Hadoop	Suppléments	3.3.6. amzn.4-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Flink	Principal	hadoop-yarn-nodemanager	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	hadoop-ya rn-resour cemanager	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	hadoop-yarn-timelineserver	Suppléments	3.3.6. amzn.4-1. amzn2023
Flink	Principal	hudi	Suppléments	0.14.1.am zn.1-1.am zn2023
Flink	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
Flink	Principal	hadoop	Suppléments	3.3.6. amzn.4-1. amzn2023
HBase	Principal	s3-dist-cp	Suppléments	2.32.0-1 amzn2023
HCatalog	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.4-1. amzn2023
HCatalog	Principal	mariadb-c onnector-java	Suppléments	2.7.2-1
Hive	Principal	hadoop-httpfs	Suppléments	3.3.6. amzn.4-1. amzn2023
Hive	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.4-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Hive	Principal	hive-hbase	Suppléments	3.1.3. amzn.11-1 . amzn2023
Hive	Principal	mariadb-c onnector-java	Suppléments	2.7.2-1
Hive	Principal	s3-dist-cp	Suppléments	2.32.0-1 amzn2023
Phoenix	Principal	s3-dist-cp	Suppléments	2.32.0-1 amzn2023
Spark	Principal	noyau de données Spark	Suppléments	3.5.1. amzn.0-1. amzn2023
Tez	Principal	hadoop-ma preduce	Suppléments	3.3.6. amzn.4-1. amzn2023
Tez	Principal	hadoop-ma preduce-h istoryserver	Suppléments	3.3.6. amzn.4-1. amzn2023
Tez	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	bigtop-utils	Suppléments	1.2.0-1. amzn2023
Trino	Principal	hadoop-client	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	hadoop-hdfs	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	hadoop-hdfs-datanode	Suppléments	3.3.6. amzn.4-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Trino	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	hadoop-hdfs-namenode	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	Hadoop KMS	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	hadoop-ma preduce	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	hadoop-ma preduce-h istoryserver	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	fil Hadoop	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	hadoop-yarn-nodemanager	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	hadoop-ya rn-resour cemanager	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	hadoop-yarn-timelineserver	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	hive	Suppléments	3.1.3. amzn.11-1 . amzn2023

Application	Type de demande	Package	Type de package	Version du package
Trino	Principal	hive-hcatalog-server	Suppléments	3.1.3. amzn.11-1. amzn2023
Trino	Principal	mariadb-connector-java	Suppléments	2.7.2-1
Trino	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023
Trino	Principal	gardien de zoo	Suppléments	3.9.1. amzn.1-1. amzn2023
Trino	Principal	zookeeper-server	Suppléments	3.9.1. amzn.1-1. amzn2023
Trino	Principal	bigtop-groovy	Suppléments	2.5.4-1. amzn2023
Trino	Principal	bigtop-jsvc	Suppléments	1.2.4-1. amzn2023
Trino	Principal	hadoop	Suppléments	3.3.6. amzn.4-1. amzn2023
Trino	Principal	hadoop-lzo	Suppléments	0.4.19-1 amzn2023
Trino	Principal	catalogue hive-h	Suppléments	3.1.3. amzn.11-1. amzn2023
Trino	Principal	hive-jdbc	Suppléments	3.1.3. amzn.11-1. amzn2023
Zookeeper	Principal	emrfs	Suppléments	2.63.0-1 amzn2023
Zookeeper	Principal	hadoop-client	Suppléments	3.3.6. amzn.4-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Zookeeper	Principal	hadoop-hdfs	Suppléments	3.3.6. amzn.4-1. amzn2023
Zookeeper	Principal	hadoop-hdfs-datanode	Suppléments	3.3.6. amzn.4-1. amzn2023
Zookeeper	Principal	hadoop-hdfs-journalnode	Suppléments	3.3.6. amzn.4-1. amzn2023
Zookeeper	Principal	hadoop-hdfs-namenode	Suppléments	3.3.6. amzn.4-1. amzn2023
Zookeeper	Principal	hadoop-httpfs	Suppléments	3.3.6. amzn.4-1. amzn2023
Zookeeper	Principal	Hadoop KMS	Suppléments	3.3.6. amzn.4-1. amzn2023
Zookeeper	Principal	hadoop-ma preduce	Suppléments	3.3.6. amzn.4-1. amzn2023
Zookeeper	Principal	hadoop-ma preduce-h istoryserver	Suppléments	3.3.6. amzn.4-1. amzn2023
Zookeeper	Principal	fil Hadoop	Suppléments	3.3.6. amzn.4-1. amzn2023
Zookeeper	Principal	hadoop-yarn-nodemanager	Suppléments	3.3.6. amzn.4-1. amzn2023
Zookeeper	Principal	hadoop-yarn-proxyserver	Suppléments	3.3.6. amzn.4-1. amzn2023

Application	Type de demande	Package	Type de package	Version du package
Zookeeper	Principal	hadoop-yarn-resourcemanager	Suppléments	3.3.6. amzn.4-1. amzn2023
Zookeeper	Principal	hadoop-yarn-timelineserver	Suppléments	3.3.6. amzn.4-1. amzn2023
Zookeeper	Principal	ranger-km	Suppléments	2.0.0-1 mai 2023

notes de publication de la version 7.2.0

Les notes de mise à jour suivantes incluent des informations relatives à la version 7.2.0 d'Amazon EMR. Les modifications concernent la version 7.2.0.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 7.2.0 application upgrades include Iceberg 1.5.0-amzn-0 and Delta 3.1.0.
- Amazon EMR ajoute un support qui vous permet d'utiliser d'autres applications telles que HBase Flink et Hive avec la classe de stockage Amazon S3 Express One Zone.
- Cette version ajoute la capacité de lire les objets restaurés, afin que vous puissiez lire les objets Glacier depuis un emplacement S3 avec le `S3A` protocole. Cette fonctionnalité fonctionne avec Spark, Flink et Hive.
- Étiquettes de nœuds avec dimensionnement géré : utilisez des étiquettes de nœuds avec vos instances en fonction du type de marché ou du type de nœud pour améliorer le dimensionnement automatique à partir d'Amazon EMR. Pour plus d'informations, consultez la section Utilisation du dimensionnement géré dans Amazon EMR.

Problèmes connus

- Python 3.11 n'est pas pris en charge avec EMR Studio.

Modifications, améliorations et problèmes résolus

- Cette version corrige un problème de blocage qui peut survenir lors des opérations internes de nettoyage des étapes. Cette opération gère le cycle de vie des étapes au fur et à mesure qu'elles se terminent sur le cluster EMR. Ce problème concerne les opérations critiques d'Amazon EMR, telles que le fonctionnement des étapes et le dimensionnement.
- Cette version résout un problème AMIs selon lequel les clusters personnalisés dotés de certains fichiers journaux préexistants peuvent provoquer l'échec du démon de gestion des journaux Amazon EMR.
- Amazon EMR 7.2.0 met à niveau le daemon Amazon EMR responsable de la gestion des clusters et des activités de surveillance du SDK v1 vers v2. AWS

- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250527.1	6.1.134-152.225 mai 2023	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250512,(	6.1.134-152.225 mai 2023	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Bahreïn) (Émirats arabes unis), Canada (centre), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			, AWS GovCloud (États-Unis est), AWS GovCloud (États-Unis ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250428.1	6,1134-150,224 mai 2023	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Malaisie), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250414,(	6.1.132-147.221 mai 2023	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7.2 250331.(	6.1.131-143.221 mai 2023	18 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre), Asie-Pacifique (Tokyo)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250303.(	6.1.129-138.220 août 2023	27 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250218,:	6.1.128-136.201. mai 2023	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250211.(	6.1.127-135.201. mai 2023	26 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250123.4	6.1.124-134.200. mai 2023	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250115,(	6.1.119-129.201.am zn2023	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 241121.(	6.1.115-126.197. mai 2023	12 décembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 241031.(	6.1.112-124.190. mai 2023	15 novembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,52(41001.1	6 janvier 109 à 118,189 mai 2023	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5,2 240819,(	6.1.102-111.182 mai 2023	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5,2 240730,(	6,197-104,177 mai 2023	2 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5,2 240722,(	6,197-104,177 mai 2023	24 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,52(40708.0	6,196-102,177 mai 2023	8 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv), Canada Ouest (Calgary) AWS

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			GovCloud , (États-Unis Ouest), (États-Unis Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

Versions Java 7.2.0 par défaut

Les versions 7.2 et supérieures d'Amazon EMR sont fournies par défaut avec Amazon Corretto 17 (basé sur OpenJDK) pour les applications compatibles avec Corretto 17 (JDK 17), à l'exception d'Apache Livy.

Le tableau suivant indique les versions Java par défaut pour les applications dans Amazon EMR 7.2.0. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	21, 17 ANS
Zeppelin	8
Pig	8
Zookeeper	8

Versions des composants 7.2.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
adot-java-agent	1,31,0	Agent Java qui collecte des métriques à partir des démons d'application.
delta	3.1.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-amazon-cloudwatch-agent	1,300032,2-amzn-0	Une application qui collecte des métriques internes au niveau du système et des métriques d'applica tion personnalisées à partir d'instances Amazon EC2 .
emr-ddb	5.4.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.11.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3,15,0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.16.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.32.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.11.0	Connecteur S3Select EMR
emr-wal-cli	1.2.2	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,63,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1,18.1-amzn-1	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1,18.1-amzn-1	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
hadoop-client	3.3.6-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.3.6-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.6-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.3.6-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	3.3.6-amzn-4	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.6-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.6-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.6-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.3.6-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.3.6-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2,4,17-amzn-6	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	2,4,17-amzn-6	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2,4,17-amzn-6	HBase client en ligne de commande.
hbase-rest-server	2,4,17-amzn-6	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2,4,17-amzn-6	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2,4,17-amzn-6	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-11	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-11	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-11	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-11	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-11	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	3.1.3-amzn-11	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-11	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0,14.1-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0,14.1-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0,14.1-amzn-1	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0,14.1-amzn-1	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1,5.0-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.

Composant	Version	Description
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.8.0 - incubation	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,8.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.3	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	5.1.3	Connecteurs Apache Phoenix pour Spark-3

Composant	Version	Description
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0,285-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0,285-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0,285-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	436-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	436-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	436-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.

Composant	Version	Description
r	4.3.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.5.1-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.5.1-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.5.1-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.5.1-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	24.02.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-9	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
tez-on-worker	0.10.2-amzn-9	Les bibliothèques et l'application tez YARN pour les nœuds de travail.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.9.1-amzn-1	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.9.1-amzn-1	ZooKeeper client en ligne de commande.

Classifications de configuration 7.2.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

classifications EMR-7.2.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.

Classifications	Description	Actions de reconfiguration
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop	Not available.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Classifications	Description	Actions de reconfiguration
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Classifications	Description	Actions de reconfiguration
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.

Classifications	Description	Actions de reconfiguration
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Modifiez les valeurs dans ZooKeeper le fichier logback.x ml.	Restarts Zookeeper server.
emr-metrics	Modifiez les paramètres métriques EMR pour ce nœud.	Restarts the CloudWatc hAgent service.

journal des modifications 7.2.0

Journal des modifications pour la version 7.2.0 et notes de version

Date	Événement	Description
25/07/2024	Publication de documents	Notes de mise à jour d'Amazon EMR 7.2.0 publiées pour la première fois
16/07/2024	Première version	Amazon EMR 7.2.0 a été déployé pour la première fois dans les régions commercia les initiales

Amazon EMR version 7.1.0

Versions de l'application 7.1.0

Cette version inclut les applications suivantes : AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, OoziePhoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS SDK pour Java	2,25,70, 1,12,747	2,23,18, 1,12,705	2,23,18, 1,12,656	2.20.160-amzn-0, 1.12.569
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3.9
Scala	2,1,18	2,1,17	2,1,17	2,1,17
AmazonCloudWatchAgent	1,300032,2	1,300032,2	1,300032,2	1,300031,1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18.1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4,17	2.4,17	2.4,17	2.4,17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
Hudi	0,15,0-amzn-0	0,14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1,5.0-amzn-0	1.4.3-amzn-0	1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,285	0,285	0,284	0,283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

notes de publication de la version 7.1.0

Les notes de mise à jour suivantes incluent des informations relatives à la version 7.1.0 d'Amazon EMR. Les modifications concernent la version 7.1.0.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 7.1.0 application upgrades include Livy 0.8.0, Trino 435, and ZooKeeper 3.9.1.
- Remplacement de nœuds défectueux — Avec Amazon EMR 7.1.0 et versions ultérieures, le remplacement de nœuds défectueux est activé par défaut. Amazon EMR remplacera donc facilement vos nœuds défectueux. Pour éviter d'affecter vos flux de travail existants sur les versions 7.0.0 et antérieures d'Amazon EMR, le remplacement des nœuds défectueux est désactivé si vous avez activé la protection contre la résiliation dans votre cluster.
- CloudWatch Agent : configurez l' CloudWatch agent pour utiliser des métriques système supplémentaires, ajouter des métriques d'application et modifier la destination des métriques avec l'API de configuration Amazon EMR.

Problèmes connus

- L'application JupyterEnterpriseGateway (JEG) ne prend pas en charge Python 3.11 dans cette version.

Modifications, améliorations et problèmes résolus

- Alors qu'Amazon EMR 7.1.0 prend en charge Python 3.9 par défaut, Livy 0.8.0 et Spark dans Amazon EMR 7.1.0 prennent en charge Python 3.11.
- Cette version corrige le problème de devoir exécuter chaque ligne une par une lors de l'utilisation PySpark avec Python version 3.11.
- Mise à niveau de Zeppelin — Amazon EMR 7.1.0 inclut une mise à niveau de Zeppelin vers le SDK pour Java v2. AWS Cette mise à niveau permet à un ordinateur portable Zeppelin S3 d'accepter un fournisseur de matériel de chiffrement personnalisé. Le AWS SDK pour Java v2 supprime EncryptionMaterialsProvider l'interface. Lorsque vous effectuez une mise à niveau vers Amazon EMR 7.1.0, vous devez implémenter l'interface Keyring si vous souhaitez utiliser un chiffrement personnalisé. Pour un exemple de mise en œuvre de l'interface Keyring, consultez KmsKeyring.java.

- Lors de la mise à niveau vers la version 7.1.0 d'Amazon EMR, modifiez votre fournisseur de clés personnalisé pour le chiffrement du disque local afin de générer des clés à l'aide de l'algorithme AES de. AES/GCM/NoPadding Si vous ne mettez pas à jour l'algorithme, la création du cluster risque d'échouer avec le message d'erreur. `Local disk encryption failed on master instance (i-123456789) due to internal error.` Pour plus d'informations sur la création d'un fournisseur de clés personnalisé, voir Création d'un fournisseur de clés personnalisé.
- Amazon EMR 7.1.0 améliore la résilience d'un nœud dans des conditions d'espace disque restreint en améliorant la logique de troncature des journaux pour les fichiers dont les descripteurs de fichier sont ouverts.
- Cette version améliore la logique de codage et de décodage afin de minimiser le risque de corruption des données et de défaillance du nœud, les démons Amazon EMR lisant et écrivant des fichiers lors du redémarrage d'un nœud.

- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250527.1	6.1.134-152.225 mai 2023	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250512,(	6.1.134-152.225 mai 2023	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Bahreïn) (Émirats arabes unis), Canada (centre), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			, AWS GovCloud (États-Unis est), AWS GovCloud (États-Unis ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250428.1	6,1134-150,224 mai 2023	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourne), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Malaisie), Asie-Pacifique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250414,(	6.1.132-147.221 mai 2023	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7.2 250331.(	6.1.131-143.221 mai 2023	18 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre), Asie-Pacifique (Tokyo)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250303.(	6.1.129-138.220 août 2023	27 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250218,:	6.1.128-136.201. mai 2023	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250211.(	6.1.127-135.201. mai 2023	26 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250123.4	6.1.124-134.200. mai 2023	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250115,(	6.1.119-129.201.am zn2023	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 241121.(	6.1.115-126.197. mai 2023	12 décembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 241031.(	6.1.112-124.190. mai 2023	15 novembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,52(41001.1	6 janvier 109 à 118,189 mai 2023	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5,2 240819,(	6.1.102-111.182 mai 2023	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5,2 240730,(	6,197-104,177 mai 2023	2 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5,2 240722,(	6,197-104,177 mai 2023	24 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,52(40708,0	6,196-102,177 mai 2023	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,3.2 240219,(	6.1.77-99,164 mai 2023	8 mai 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv), Canada Ouest (Calgary) AWS

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			GovCloud , (États-Unis Ouest), (États-Unis Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

Versions Java 7.1.0 par défaut

Amazon EMR version 7.1 et supérieure est livré avec Amazon Corretto 17 (basé sur OpenJDK) par défaut pour les applications qui prennent en charge Corretto 17 (JDK 17), à l'exception d'Apache Livy.

Le tableau suivant indique les versions Java par défaut pour les applications dans Amazon EMR 7.1.0. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

Versions des composants 7.1.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
adot-java-agent	1,31,0	Agent Java qui collecte des métriques à partir des démons d'application.
delta	3.0.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-amazon-cloudwatch-agent	1,300032,2-amzn-0	Une application qui collecte des métriques internes au niveau du système et des métriques d'applica tion personnalisées à partir d'instances Amazon EC2 .
emr-ddb	5.3.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.10.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3,14,0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.31.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.10.0	Connecteur S3Select EMR
emr-wal-cli	1.2.2	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,62,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1,18.1-amzn-0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1,18.1-amzn-0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
hadoop-client	3.3.6-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.3.6-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.6-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.3.6-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	3.3.6-amzn-3	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.6-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.6-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.6-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.3.6-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.3.6-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2,4,17-amzn-5	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	2,4,17-amzn-5	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2,4,17-amzn-5	HBase client en ligne de commande.
hbase-rest-server	2,4,17-amzn-5	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2,4,17-amzn-5	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2,4,17-amzn-5	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-10	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-10	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-10	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-10	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-10	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	3.1.3-amzn-10	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-10	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.14.1-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.14.1-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.14.1-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.14.1-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1.4.3-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.

Composant	Version	Description
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.8.0 - incubation	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,8.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.3	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	5.1.3	Connecteurs Apache Phoenix pour Spark-3

Composant	Version	Description
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0,284-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0,284-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0,284-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	435-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	435-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	435-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.

Composant	Version	Description
r	4.3.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.5.0-amzn-1	Clients de ligne de commande Spark.
spark-history-server	3.5.0-amzn-1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.5.0-amzn-1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.5.0-amzn-1	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	23.10.0-amzn-1	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-8	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
tez-on-worker	0.10.2-amzn-8	Les bibliothèques et l'application tez YARN pour les nœuds de travail.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.9.1-amzn-0	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.9.1-amzn-0	ZooKeeper client en ligne de commande.

Classifications de configuration 7.1.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

classifications EMR-7.1.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.

Classifications	Description
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
docker-conf	Modifiez les paramètres liés au docker.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-session	Modifiez les paramètres Flink log4j-ses sion.properties pour la session. Kubernetes/ Yarn
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive	Paramètres définis par Amazon EMR pour Apache Hive.

Classifications	Description
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.

Classifications	Description
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defa ults.conf de Hudi.
iceberg-defaults	Modification des valeurs du fichier iceberg-d efaults.conf d'Iceberg.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-note s Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.

Classifications	Description
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.

Classifications	Description
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Trino.
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.

Classifications	Description
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.properties de Trino.
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.

Classifications	Description
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.

Classifications	Description
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-logback	Modifiez les valeurs dans ZooKeeper le fichier logback.xml.

Classifications	Description
emr-metrics	Modifiez les paramètres métriques EMR pour ce nœud.

journal des modifications 7.1.0

Journal des modifications pour la version 7.1.0 et notes de version

Date	Événement	Description
2024-05-10	Publication de documents	Notes de mise à jour d'Amazon EMR 7.1.0 publiées pour la première fois
08-05-2024/	Déploiement terminé	Amazon EMR 7.1.0 entièrement déployé dans toutes les régions prises en charge
23/04/2024	Première version	Amazon EMR 7.1.0 a été déployé pour la première fois dans les régions commercia les initiales

Amazon EMR version 7.0.0

Versions d’application 7.0.0

Cette version inclut les applications suivantes : AmazonCloudWatchAgentDeltaFlink, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, OoziePhoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS SDK pour Java	2,25,70, 1,12,747	2,23,18, 1,12,705	2,23,18, 1,12,656	2.20.160-amzn-0, 1.12.569
Python	3,9, 3,11	3,9, 3,11	3,9, 3,11	3.9
Scala	2,1,18	2,1,17	2,1,17	2,1,17
AmazonCloudWatchAgent	1,300032,2	1,300032,2	1,300032,2	1,300031,1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18.1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4,17	2.4,17	2.4,17	2.4,17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0,15,0-amzn-0	0,14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1,5.0-amzn-0	1.4.3-amzn-0	1.4.2-amzn-0

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,285	0,285	0,284	0,283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

Notes de mise à jour 7.0.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 7.0.0. Les modifications ont été apportées à la version 6.15.0.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 7.0.0 application upgrades include Python 3.9, Spark 3.5, Flink 1.18, and Delta 3.0. Cette version ajoute également la prise en charge de l'application CloudWatch Agent Amazon et abandonne la prise en charge de Ganglia.
- Amazon Corretto 17 – Les versions 7.0 et ultérieures d'Amazon EMR sont fournies par défaut avec Amazon Corretto 17 (basé sur OpenJDK) pour les applications qui prennent en charge Corretto 17 (JDK 17), à l'exception d'Apache Livy. Pour plus d'informations sur les versions de JDK prises en charge pour les applications de cette version, voir la rubrique Versions Java par défaut 7.0.0.
- Amazon Linux 2023 — Avec la version 7.0, les clusters Amazon EMR s'exécutent désormais sur AL2 023 par défaut. Pour savoir comment cela affecte la version d'AMI par défaut, voir la rubrique Considérations relatives aux mises à jour du Guide de gestion Amazon EMR. Notez également que AL2 023 a supprimé Python 2.7, de sorte que tous les composants nécessitant Python doivent désormais être écrits avec Python 3.
- S3 sur Outposts avec **`s3a`** – Amazon EMR prend désormais en charge les compartiments Amazon S3 sur Outposts avec le système de fichiers `s3a`. Pour plus d'informations sur S3 sur Outposts, voir la rubrique Qu'est-ce que S3 sur Outposts ? du Guide de l'utilisateur Amazon Simple Storage Service.

Problèmes connus

- Notez que vous ne pouvez pas exécuter plus d'une ligne à la fois si vous utilisez PySpark Python 3.10 ou version ultérieure. Vous devez exécuter chaque ligne une par une.

Modifications, améliorations et problèmes résolus

- Pour maintenir l'état de toutes les instances d'un cluster, Amazon EMR s'intègre à Apache YARN, Apache HDFS et Kerberos. Avec la version 7.0, nous avons amélioré ces intégrations afin de supprimer en toute sécurité l'état des instances arrêtées en raison d'une mise à l'échelle ou d'autres opérations. Cette amélioration est particulièrement importante pour les clusters longue durée pour lesquels la mise à l'échelle gérée est activée, car font régulièrement l'objet de mises à l'échelle et accumulent des milliers d'instances arrêtées au cours de leur durée de vie.
- Cette version améliore la configuration de Kerberos afin d'inclure uniquement la prise en charge des chiffrements basés sur AES. Le KDC Kerberos avec des chiffrements non basés sur AES n'est plus pris en charge avec les clusters EMR qui s'exécutent sur la version 7.0.0 et les versions

ultérieures d'Amazon EMR. Les chiffrements basés sur AES offrent la meilleure sécurité pour vos clusters.

- Dans le cadre de la migration vers le AWS SDK 2.x, Amazon EMR 7.0 inclut une mise à jour du connecteur Spark Kinesis pour des raisons de compatibilité. Cette mise à jour n'est pas disponible dans la version communautaire d'Apache Spark. Si vous utilisez le connecteur Spark Kinesis sur une version d'Amazon EMR antérieure à la 7.0, vous devez migrer les codes de vos applications pour qu'ils s'exécutent sur SDK 2.x avant de migrer vos charges de travail vers Amazon EMR 7.0. Pour de plus amples informations, veuillez consulter Migration du connecteur Spark Kinesis vers le SDK 2.x pour Amazon EMR 7.0.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250527.1	6.1.134-152.225 mai 2023	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon) AWS GovCloud , ouest des États-Unis (N. Californie), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (Centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250512,(	6.1.134-152.225 mai 2023	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Bahreïn) (Émirats arabes unis), Canada (centre), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			, AWS GovCloud (États-Unis est), AWS GovCloud (États-Unis ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250428.1	6,1134-150,224 mai 2023	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Bahreïn) (Émirats arabes unis), Canada (centre), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			, AWS GovCloud (États-Unis est), AWS GovCloud (États-Unis ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7,2 250414,(	6.1.132-147.221 mai 2023	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7.2 250331.(	6.1.131-143.221 mai 2023	18 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Pacifique (Tokyo)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250303.(	6.1.129-138.220 août 2023	27 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250218,:	6.1.128-136.201. mai 2023	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250211.(	6.1.127-135.201. mai 2023	26 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250123.4	6.1.124-134.200. mai 2023	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250115,(	6.1.119-129.201.am zn2023	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 241121.(	6.1.115-126.197. mai 2023	12 décembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 241031.(	6.1.112-124.190. mai 2023	15 novembre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,52(41001.1	6 janvier 109 à 118,189 mai 2023	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5,2 240819,(	6.1.102-111.182 mai 2023	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5,2 240730,(	6,197-104,177 mai 2023	2 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,5,2 240722,(	6,197-104,177 mai 2023	24 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,52(40708,0	6,196-102,177 mai 2023	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv), Canada Ouest (Calgary), (US-Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (US-Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,3.2 240304,(	6.1.79-99,164 mai 2023	12 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv), Canada Ouest (Calgary) AWS

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			GovCloud , (États-Unis Ouest), (États-Unis Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,3.2 240219,(	6.1.77-99,164 mai 2023	1er mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv), Canada Ouest (Calgary) AWS

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			GovCloud , (États-Unis Ouest), (États-Unis Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,3.2 240205,(	6,175-99,163 mai 2023	19 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv), Canada Ouest (Calgary) AWS

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			GovCloud , (États-Unis Ouest), (États-Unis Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,3.2 240122.(	6,172-96,166 mai 2023	5 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv), Canada Ouest (Calgary) AWS

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			GovCloud , (États-Unis Ouest), (États-Unis Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,3.2 240108.(	6,172-96,166 mai 2023	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv), Canada Ouest (Calgary) AWS

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			GovCloud , (États-Unis Ouest), (États-Unis Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.3.2 23 1211,4	6.1.66-91.160.amzn2023	19 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

Versions Java par défaut 7.0.0

La version 7.0 et les versions ultérieures d'Amazon EMR sont fournies par défaut avec Amazon Corretto 17 (basé sur OpenJDK) pour les applications qui prennent en charge Corretto 17 (JDK 17), à l'exception d'Apache Livy.

Le tableau suivant présente les versions de Java par défaut pour les applications dans Amazon EMR 7.0.0 : Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	11, 8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

Versions des composants 7.0.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
adot-java-agent	1,31,0	Agent Java qui collecte des métriques à partir des démons d'application.
delta	3.0.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
delta-standalone-connectors	0.6.0	Les connecteurs Delta fournissent différents environnements d'exécution pour intégrer Delta Lake avec des moteurs tels que Flink, Hive et Presto.
emr-amazon-cloudwatch-agent	1.300031.1-amzn-0	Une application qui collecte des métriques internes au niveau du système et des métriques d'applica tion personnalisées à partir d'instances Amazon EC2 .
emr-ddb	5.2.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.9.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.

Composant	Version	Description
emr-kinesis	3,13,0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.30.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.9.0	Connecteur S3Select EMR
emr-wal-cli	1.2.1	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,61,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.18.0-amzn-0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.18.0-amzn-0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
hadoop-client	3.3.6-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.3.6-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.6-amzn-2	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	3.3.6-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.3.6-amzn-2	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.6-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.6-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.6-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.3.6-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.3.6-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	2.4.17-amzn-4	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.17-amzn-4	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.17-amzn-4	HBase client en ligne de commande.
hbase-rest-server	2.4.17-amzn-4	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.17-amzn-4	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.17-amzn-4	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-9	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-9	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-9	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	3.1.3-amzn-9	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-9	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-9	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-9	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.14.0-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.14.0-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.14.0-amzn-1	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.14.0-amzn-1	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
iceberg	1.4.2-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,8.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.3	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-connectors	5.1.3	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.283-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.283-amzn-1	Service pour exécuter les éléments d'une requête.
presto-client	0.283-amzn-1	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	426-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	426-amzn-1	Service pour exécuter les éléments d'une requête.
trino-client	426-amzn-1	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.

Composant	Version	Description
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.3.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.5.0-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.5.0-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.5.0-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.5.0-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	23.10.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.

Composant	Version	Description
tez-on-yarn	0.10.2-amzn-7	Les bibliothèques et l'application tez YARN.
tez-on-worker	0.10.2-amzn-7	Les bibliothèques et l'application tez YARN pour les nœuds de travail.
webserver	2,4,58	Serveur HTTP Apache.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 7.0.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-7.0.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.

Classifications	Description
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
docker-conf	Modifiez les paramètres liés au docker.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase

Classifications	Description
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.

Classifications	Description
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive	Paramètres définis par Amazon EMR pour Apache Hive.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.

Classifications	Description
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.

Classifications	Description
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Trino.
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.

Classifications	Description
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.properties de Trino.
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.

Classifications	Description
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
spark	Paramètres définis par Amazon EMR pour Apache Spark.

Classifications	Description
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.

Classifications	Description
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Journal des modifications 7.0.0

Journal des modifications pour la version 7.0.0 et les notes de mise à jour

Date	Événement	Description
29/12/2023	Déploiement terminé	Amazon EMR 7.0.0 entièrement déployé dans toutes les régions prises en charge
28/12/2023	Publication de documents	Notes de mise à jour d'Amazon EMR 7.0.0 publiées pour la première fois
19/12/2023	Première version	Amazon EMR 7.0.0 a été déployé pour la première fois dans des régions commercia les limitées

Versions Amazon EMR 6.x

Cette section contient les versions des applications, les notes de mise à jour, les versions des composants et les classifications de configuration disponibles dans chaque version 6.x d'Amazon EMR.

Lorsque vous lancez un cluster, vous pouvez choisir parmi plusieurs versions d'Amazon EMR. Vous pouvez ainsi tester et utiliser des versions d'applications qui répondent à vos exigences de compatibilité. Vous spécifiez le numéro de version avec l'étiquette de version. Les étiquettes de version sont sous la forme `emr-x.x.x`. Par exemple, `emr-7.9.0`.

Les nouvelles versions d'Amazon EMR sont mises à disposition dans différentes régions sur une période de plusieurs jours, en commençant par la première région à la date de sortie initiale. Il est possible que la dernière version ne soit pas disponible dans votre région pendant cette période.

Pour un tableau complet des versions des applications dans chaque version 6.x d'Amazon EMR, consultez Versions des applications dans les versions 6.x d'Amazon EMR.

Rubriques

- Versions des applications dans les versions 6.x d'Amazon EMR
- Amazon EMR version 6.15.0
- Amazon EMR version 6.14.0
- Amazon EMR version 6.13.0
- Amazon EMR version 6.12.0
- Amazon EMR version 6.11.1
- Amazon EMR version 6.11.0
- Amazon EMR version 6.10.1
- Amazon EMR version 6.10.0
- Amazon EMR version 6.9.1
- Amazon EMR version 6.9.0
- Amazon EMR version 6.8.1
- Amazon EMR version 6.8.0
- Amazon EMR version 6.7.0
- Amazon EMR version 6.6.0
- Amazon EMR version 6.5.0
- Amazon EMR version 6.4.0
- Amazon EMR version 6.3.1
- Amazon EMR version 6.3.0
- Amazon EMR version 6.2.1
- Amazon EMR version 6.2.0
- Amazon EMR version 6.1.1
- Amazon EMR version 6.1.0
- Amazon EMR version 6.0.1

- Amazon EMR version 6.0.0

Versions des applications dans les versions 6.x d'Amazon EMR

Pour obtenir un tableau complet répertoriant les versions d'applications disponibles dans chaque version d'Amazon EMR 6.x, ouvrez Versions des applications dans les versions 6.x d'Amazon EMR dans votre navigateur.

Amazon EMR version 6.15.0

Versions d’application 6.15.0

Cette version inclut les applications suivantes : Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
AWS SDK pour Java	2.20.160-amzn-0, 1.12.569	1,12,543	1,12,513	1,12,490
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,17	2,1,15	2,1,15	2,1,15

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.4.0	2.4.0
Flink	1,17.1	1,17.1	1.17.0	1.17.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4,17	2.4,17	2.4,17	2.4,17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.14.0-amzn-0	0.13.1-amzn-2	0.13.1-amzn-1	0.13.1-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.4.0-amzn-0	1.3.1-amzn-0	1.3.0-amzn-1	1.3.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
Presto	0,283	0,281	0,281	0,281
Spark	3.4.1	3.4.1	3.4.1	3.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	426	422	414	414
Zeppelin	0,1,1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notes de mise à jour 6.15.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.15.0. Les modifications ont été apportées à la version 6.14.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.15.0.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 6.15.0 application upgrades include Apache Hadoop 3.3.6, Apache Hudi 0.14.0-amzn-0, Iceberg 1.4.0-amzn-0, and Trino 426..
- Lancement plus rapide pour les clusters EMR exécutés sur EC2 — Le lancement d'un cluster Amazon EMR sur un cluster est désormais jusqu'à 35 % plus rapide. EC2 Grâce à cette amélioration, la plupart des clients peuvent lancer leurs clusters en moins de 5 minutes.
- CodeWhisperer pour EMR Studio — Vous pouvez désormais utiliser Amazon avec CodeWhisperer Amazon EMR Studio pour obtenir des recommandations en temps réel lorsque vous écrivez du code. JupyterLab CodeWhisperer peut compléter vos commentaires, terminer des lignes de code individuelles, faire line-by-line des recommandations et générer des fonctions entièrement formées.
- Redémarrage accéléré des tâches avec Flink – Avec Amazon EMR 6.15.0 et versions ultérieures, plusieurs nouveaux mécanismes sont disponibles pour Apache Flink afin d'améliorer le temps de

redémarrage des tâches lors des opérations de récupération ou de mise à l'échelle des tâches. Cela permet d'optimiser la vitesse de récupération et de redémarrage des graphes d'exécution afin d'améliorer la stabilité des tâches.

- Contrôle d'accès détaillé et au niveau des tables pour les formats de tables ouvertes : avec Amazon EMR 6.15.0 et versions ultérieures, lorsque vous exécutez des tâches Spark sur Amazon EMR sur des clusters EC2 qui accèdent aux données du catalogue de données AWS Glue, vous pouvez appliquer des autorisations au niveau des tables, des lignes, des colonnes et des cellules AWS Lake Formation aux tables basées sur Hudi, Iceberg ou Delta Lake.
- Mise à niveau de Hadoop – Amazon EMR 6.15.0 inclut une mise à niveau d'Apache Hadoop vers la version 3.3.6. Hadoop 3.3.6 était la dernière version au moment du déploiement d'Amazon EMR 6.15, publiée par Apache en juin 2023. Les versions précédentes d'Amazon EMR (6.9.0 à 6.14.x) utilisaient Hadoop 3.3.3.

  La mise à niveau inclut des centaines d'améliorations et de correctifs, ainsi que des fonctionnalités telles que des paramètres de nœud de données reconfigurables, l'option `DFSAdmin` qui permet de lancer des opérations de reconfiguration en bloc sur tous les nœuds de données actifs et une API vectorielle qui permet aux lecteurs à forte densité de recherche de spécifier plusieurs plages à lire. Hadoop 3.3.6 ajoute également la prise en charge du HDFS APIs et de la sémantique à son journal d'écriture anticipée (WAL), afin qu' HBase il puisse fonctionner sur d'autres implémentations de systèmes de stockage. Pour plus d'informations, consultez les journaux des modifications pour les versions 3.3.4, 3.3.5 et 3.3.6 dans la documentation Apache Hadoop.
- Support du AWS SDK pour Java, version 2 - Les applications Amazon EMR 6.15.0 peuvent utiliser les versions 1.12.569 ou 2.20.160 du SDK pour AWS Java si l'application prend en charge la version v2. Le AWS SDK pour Java 2.x est une réécriture majeure de la base de code de la version 1.x. Il repose sur Java 8+ et ajoute plusieurs fonctionnalités fréquemment demandées. Cela inclut la prise en charge d'E/S non bloquantes et la possibilité de brancher une implémentation HTTP différente au moment de l'exécution. Pour plus d'informations, notamment un Guide de migration du kit SDK pour la version 1 de Java version la version 2, voir le guide du Kit AWS SDK pour Java, version 2.

Modifications, améliorations et problèmes résolus

- Pour améliorer vos clusters EMR haute disponibilité, cette version assure la connectivité aux démons Amazon EMR sur un hôte local utilisant des points de terminaison IPv6.
- Cette version permet au protocole TLS 1.2 de communiquer avec les nœuds ZooKeeper provisionnés sur tous les nœuds principaux de votre cluster à haute disponibilité.

- Cette version améliore la gestion des fichiers journaux de ZooKeeper transactions conservés sur les nœuds principaux afin de minimiser les scénarios dans lesquels les fichiers journaux dépassent les limites et interrompent les opérations du cluster.
- Cette version accentue la résilience de la communication intra-nœud pour les clusters EMR haute disponibilité. Cette amélioration réduit le risque d'échec des actions d'amorçage ou de démarrage du cluster.
- Tez dans Amazon EMR 6.15.0 introduit des configurations que vous pouvez spécifier pour ouvrir de manière asynchrone les divisions d'entrée dans une division groupée Tez. Cela permet d'accélérer les performances des requêtes de lecture lorsqu'il existe un grand nombre de divisions d'entrée dans une seule division groupée Tez. Pour plus d'informations, consultez la rubrique Ouverture fractionnée asynchrone Tez.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique (Taipei), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Jakarta) (Malaisie), Asie-Pacifique (Melbourne), Asie-Paci fique (Mumbai), Asie-Paci fique (Osaka), Asie-Paci fique (Séoul), Asie-Paci fique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Malaisie), Asie-Pacifique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Émirats arabes unis), Canada (centre),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Canada Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Calgary), Europe (Zurich), Asie-Pacifique (Malaisie)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	13 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

Versions Java par défaut 6.15.0

La version 6.12.0 et les versions ultérieures d'Amazon EMR prennent en charge toutes les applications dotées d'Amazon Corretto 8 par défaut, à l'exception de Trino. Pour Trino, Amazon EMR prend en charge Amazon Corretto 17 par défaut à partir de la version 6.9.0 d'Amazon EMR. Amazon EMR prend également en charge certaines applications avec Amazon Corretto 11 et 17. Ces applications sont répertoriées dans le tableau suivant. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Bien qu'Amazon EMR prenne en charge Amazon Corretto 11 et 17 sur Apache Spark, Apache Hadoop et Apache Hive, les performances peuvent régresser pour certaines charges de travail lorsque vous utilisez ces versions de Corretto. Nous vous recommandons de tester vos charges de travail avant de modifier les valeurs par défaut.

Le tableau suivant présente les versions de Java par défaut pour les applications dans Amazon EMR 6.15.0 :

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

Versions des composants 6.15.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.2	Kit de développement logiciel Amazon SageMaker Spark
delta	2.4.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
delta-standalone-connectors	0.6.0	Les connecteurs Delta fournissent différents environnements d'exécution pour intégrer Delta Lake avec des moteurs tels que Flink, Hive et Presto.
emr-ddb	5.2.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.

Composant	Version	Description
emr-goodies	3.8.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.12.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.29.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.8.0	Connecteur S3Select EMR
emr-wal-cli	1.2.0	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,60,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.17.1-amzn-1	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.17.1-amzn-1	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.3.6-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.3.6-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.6-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.3.6-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.3.6-amzn-1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.6-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.6-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	3.3.6-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.6-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.3.6-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.3.6-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.17-amzn-3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.17-amzn-3	Service pour desservir une ou plusieurs HBase régions.
hbase-client	2.4.17-amzn-3	HBase client en ligne de commande.
hbase-rest-server	2.4.17-amzn-3	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.17-amzn-3	Service fournissant un point de terminaison Thrift pour. HBase

Composant	Version	Description
hbase-operator-tools	2.4.17-amzn-3	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-8	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-8	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-8	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-8	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-8	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-8	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-8	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.14.0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.

Composant	Version	Description
hudi-presto	0.14.0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.14.0-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.14.0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1.4.0-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.

Composant	Version	Description
nvidia-cuda	11,8.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.3	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	5.1.3	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.283-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.283-amzn-0	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
presto-client	0.283-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	426-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	426-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	426-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.4.1-amzn-2	Clients de ligne de commande Spark.
spark-history-server	3.4.1-amzn-2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	3.4.1-amzn-2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.4.1-amzn-2	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	23.08.1-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-6	Les bibliothèques et l'applica tion tez YARN.
tez-on-worker	0.10.2-amzn-6	Les bibliothèques et l'applica tion tez YARN pour les nœuds de travail.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 6.15.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.15.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classifications	Description	Actions de reconfiguration
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.

Classifications	Description	Actions de reconfiguration
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.pr operties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-s ite.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.pr operties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.prope rties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_c onfig.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistan ce S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.

Classifications	Description	Actions de reconfiguration
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.p roperties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.

Classifications	Description	Actions de reconfiguration
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Journal des modifications 6.15.0

Journal des modifications pour la version 6.15.0 et les notes de mise à jour

Date	Événement	Description
17/11/2023	Publication de documents	Notes de mise à jour d'Amazon EMR 6.15.0 publiées pour la première fois
17/11/2023	Déploiement terminé	Amazon EMR 6.15.0 entièrement déployé dans toutes les régions prises en charge
13/11/2023	Première version	Amazon EMR 6.15.0 a été déployé pour la première fois dans des régions commercia les limitées

Amazon EMR version 6.14.0

Versions d'application 6.14.0

Cette version inclut les applications suivantes : Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.14.0	emr-6.13.0	emr-6.12.0	emr-6.11.1
AWS SDK pour Java	1,12,543	1,12,513	1,12,490	1,12,446
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,15	2,1,15	2,1,15	2,1,15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.4.0	2.2.0
Flink	1,17.1	1.17.0	1.17.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4,17	2.4,17	2.4,17	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-2	0.13.1-amzn-1	0.13.1-amzn-0	0.13.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.3.1-amzn-0	1.3.0-amzn-1	1.3.0-amzn-0	1.2.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1

	emr-6.14.0	emr-6.13.0	emr-6.12.0	emr-6.11.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,281	0,281	0,281	0,279
Spark	3.4.1	3.4.1	3.4.0	3.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	422	414	414	410
Zeppelin	0,1,1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notes de mise à jour 6.14.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.14.0. Les modifications ont été apportées à la version 6.13.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.14.0.

Nouvelles fonctionnalités

- Amazon EMR 6.14.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-2, Flink 1.17.1, Iceberg 1.3.1, and Trino 422.
- La mise à l'échelle gérée par Amazon EMR est désormais disponible dans la région `ap-southeast-3` Asie-Pacifique (Jakarta) pour les clusters que vous créez avec Amazon EMR 6.14.0 ou version ultérieure.

Modifications, améliorations et problèmes résolus

- La version 6.14.0 optimise la gestion des journaux avec Amazon EMR exécuté sur Amazon. EC2 C'est pourquoi vous constaterez peut-être une légère réduction des coûts de stockage pour les journaux de votre cluster.
- La version 6.14.0 améliore le flux de travail de mise à l'échelle afin de prendre en compte les différentes instances principales dont la taille varie considérablement pour leurs volumes Amazon EBS. Cette amélioration s'applique uniquement aux nœuds principaux ; les opérations de réduction des nœuds de tâches ne sont pas affectées.
- La version 6.14.0 améliore la façon dont Amazon EMR interagit avec les applications open source telles que Apache Hadoop YARN ResourceManager and HDFS NameNode. Cette amélioration réduit le risque de retards opérationnels liés à la mise à l'échelle du cluster et atténue les échecs de démarrage dus à des problèmes de connectivité avec les applications open source.
- La version 6.14.0 optimise l'installation des applications au lancement du cluster. Cela améliore les temps de démarrage du cluster pour certaines combinaisons d'applications Amazon EMR.
- La version 6.14.0 corrige un problème où les opérations de réduction de cluster pouvaient se bloquer lorsqu'un cluster exécuté dans un VPC avec un domaine personnalisé rencontrait un redémarrage de nœud principal ou de nœud de tâche.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique (Taipei), Canada (centre), Canada Ouest (Calgary),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Europe (Espagne), Chine (Ningxia), Europe (Zurich)
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Jakarta) (Malaisie), Asie-Pacifique (Melbourne), Asie-Paci fique (Mumbai), Asie-Paci fique (Osaka), Asie-Paci fique (Séoul), Asie-Paci fique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Malaisie), Asie-Pacifique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Émirats arabes unis), Canada (centre),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Canada Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Calgary), Europe (Zurich), Asie-Pacifique (Malaisie)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	17 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	11 septembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

Versions Java par défaut 6.14.0

La version 6.12.0 et les versions ultérieures d'Amazon EMR prennent en charge toutes les applications dotées d'Amazon Corretto 8 par défaut, à l'exception de Trino. Pour Trino, Amazon EMR prend en charge Amazon Corretto 17 par défaut à partir de la version 6.9.0 d'Amazon EMR. Amazon EMR prend également en charge certaines applications avec Amazon Corretto 11 et 17. Ces applications sont répertoriées dans le tableau suivant. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Bien qu'Amazon EMR prenne en charge Amazon Corretto 11 et 17 sur Apache Spark, Apache Hadoop et Apache Hive, les performances peuvent régresser pour certaines charges de travail lorsque vous utilisez ces versions de Corretto. Nous vous recommandons de tester vos charges de travail avant de modifier les valeurs par défaut.

Le tableau suivant présente les versions de Java par défaut pour les applications dans Amazon EMR 6.14.0 :

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

Versions des composants 6.14.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.2	Kit de développement logiciel Amazon SageMaker Spark
delta	2.4.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
delta-standalone-connectors	0.6.0	Les connecteurs Delta fournissent différents environnements d'exécution pour intégrer Delta Lake avec des moteurs tels que Flink, Hive et Presto.
emr-ddb	5.1.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.7.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.11.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise

Composant	Version	Description
emr-s3-dist-cp	2.28.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.7.0	Connecteur S3Select EMR
emr-wal-cli	1.1.0	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,59,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.17.1-amzn-0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.17.1-amzn-0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	3.3.3-amzn-6	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.3.3-amzn-6	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.3-amzn-6	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.3.3-amzn-6	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.3.3-amzn-6	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.3-amzn-6	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.3-amzn-6	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-6	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.3-amzn-6	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	3.3.3-amzn-6	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.3.3-amzn-6	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.17-amzn-2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.17-amzn-2	Service pour desservir une ou plusieurs HBase régions.
hbase-client	2.4.17-amzn-2	HBase client en ligne de commande.
hbase-rest-server	2.4.17-amzn-2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.17-amzn-2	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.17-amzn-2	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-7	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	3.1.3-amzn-7	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-7	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-7	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-7	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-7	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-7	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.13.1-amzn-2	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.13.1-amzn-2	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.13.1-amzn-2	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.

Composant	Version	Description
hudi-spark	0.13.1-amzn-2	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1.3.1-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,8.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.3	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	5.1.3	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.281-amzn-2	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.281-amzn-2	Service pour exécuter les éléments d'une requête.
presto-client	0.281-amzn-2	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	422-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.

Composant	Version	Description
trino-worker	422-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	422-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.4.1-amzn-1	Clients de ligne de commande Spark.
spark-history-server	3.4.1-amzn-1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.4.1-amzn-1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.4.1-amzn-1	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	23.06.0-amzn-2	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.

Composant	Version	Description
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-5	Les bibliothèques et l'application tez YARN.
tez-on-worker	0.10.2-amzn-5	Les bibliothèques et l'application tez YARN pour les nœuds de travail.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 6.14.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.14.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop	Not available.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Classifications	Description	Actions de reconfiguration
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Classifications	Description	Actions de reconfiguration
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.

Classifications	Description	Actions de reconfiguration
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

Journal des modifications 6.14.0

Journal des modifications pour la version 6.14.0 et les notes de mise à jour

Date	Événement	Description
02/11/2023	Déploiement terminé	Amazon EMR 6.14.0 entièrement déployé dans toutes les régions prises en charge
10/2023	Publication de documents	Notes de mise à jour d'Amazon EMR 6.14.0 publiées pour la première fois
04/10/2023	Première version	Amazon EMR 6.14.0 a été déployé pour la première fois dans des régions commercia les limitées

Amazon EMR version 6.13.0

Versions d’application 6.13.0

Cette version inclut les applications suivantes : Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
AWS SDK pour Java	1,12,513	1,12,490	1,12,446	1,12,446
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,15	2,1,15	2,1,15	2,1,15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.2.0	2.2.0
Flink	1.17.0	1.17.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4,17	2.4,17	2.4.15	2.4.15

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-1	0.13.1-amzn-0	0.13.0-amzn-0	0.13.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.3.0-amzn-1	1.3.0-amzn-0	1.2.0-amzn-0	1.2.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,281	0,281	0,279	0,279
Spark	3.4.1	3.4.0	3.3.2	3.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
Trino (PrestoSQL)	414	414	410	410
Zeppelin	0,1,1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notes de mise à jour 6.13.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.13.0. Les modifications ont été apportées à la version 6.12.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.13.0.

Nouvelles fonctionnalités

- Amazon EMR 6.13.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-1, CUDA Toolkit 11.8.0, and JupyterHub 1.5.0.

Modifications, améliorations et problèmes résolus

- La version 6.13.0 améliore le démon de gestion des journaux Amazon EMR afin de garantir que tous les journaux sont chargés à une cadence régulière sur Amazon S3 lorsqu'une commande de résiliation de cluster est émise. Cela permet de résilier plus rapidement les clusters.
- La version 6.13.0 améliore les capacités de gestion des journaux Amazon EMR afin de garantir un chargement cohérent et en temps voulu de tous les fichiers journaux sur Amazon S3. Cela est particulièrement avantageux pour les clusters EMR de longue durée.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

Versions Java par défaut 6.13.0

La version 6.12.0 et les versions ultérieures d'Amazon EMR prennent en charge toutes les applications dotées d'Amazon Corretto 8 par défaut, à l'exception de Trino. Pour Trino, Amazon EMR prend en charge Amazon Corretto 17 par défaut à partir de la version 6.9.0 d'Amazon EMR. Amazon EMR prend également en charge certaines applications avec Amazon Corretto 11 et 17. Ces applications sont répertoriées dans le tableau suivant. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Bien qu'Amazon EMR prenne en charge Amazon Corretto 11 et 17 sur Apache Spark, Apache Hadoop et Apache Hive, les performances peuvent régresser pour certaines charges de travail lorsque vous utilisez ces versions de Corretto. Nous vous recommandons de tester vos charges de travail avant de modifier les valeurs par défaut.

Le tableau suivant présente les versions de Java par défaut pour les applications dans Amazon EMR 6.13.0 :

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

Versions des composants 6.13.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.2	Kit de développement logiciel Amazon SageMaker Spark
delta	2.4.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
delta-standalone-connectors	0.6.0	Les connecteurs Delta fournissent différents environnements d'exécution pour intégrer Delta Lake avec des moteurs tels que Flink, Hive et Presto.
emr-ddb	5.1.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.6.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.10.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise

Composant	Version	Description
emr-s3-dist-cp	2.27.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.6.0	Connecteur S3Select EMR
emr-wal-cli	1.1.0	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,58,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.17.0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.17.0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	3.3.3-amzn-5	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.3.3-amzn-5	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.3-amzn-5	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.3.3-amzn-5	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.3.3-amzn-5	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.3-amzn-5	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.3-amzn-5	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-5	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.3-amzn-5	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	3.3.3-amzn-5	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.3.3-amzn-5	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.17-amzn-1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.17-amzn-1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.17-amzn-1	HBase client en ligne de commande.
hbase-rest-server	2.4.17-amzn-1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.17-amzn-1	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.17-amzn-1	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-6	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	3.1.3-amzn-6	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-6	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-6	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-6	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-6	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-6	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.13.1-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.13.1-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.13.1-amzn-1	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.

Composant	Version	Description
hudi-spark	0.13.1-amzn-1	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
iceberg	1.3.0-amzn-1	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,8.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.3	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	5.1.3	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.281-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.281-amzn-1	Service pour exécuter les éléments d'une requête.
presto-client	0.281-amzn-1	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	414-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.

Composant	Version	Description
trino-worker	414-amzn-1	Service pour exécuter les éléments d'une requête.
trino-client	414-amzn-1	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.4.1-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.4.1-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.4.1-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.4.1-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	23.06.0-amzn-1	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.

Composant	Version	Description
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-4	Les bibliothèques et l'application tez YARN.
tez-on-worker	0.10.2-amzn-4	Les bibliothèques et l'application tez YARN pour les nœuds de travail.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 6.13.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.13.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop	Not available.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Classifications	Description	Actions de reconfiguration
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Classifications	Description	Actions de reconfiguration
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.

Classifications	Description	Actions de reconfiguration
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

Journal des modifications 6.13.0

Journal des modifications pour la version 6.13.0 et les notes de mise à jour

Date	Événement	Description
23/09/2023	Déploiement terminé	Amazon EMR 6.13.0 entièrement déployé dans toutes les régions prises en charge
12/09/2023	Publication de documents	Notes de mise à jour d'Amazon EMR 6.13.0 publiées pour la première fois
01/09/2023	Première version	Amazon EMR 6.13.0 a été déployé pour la première fois dans des régions commercia les limitées

Amazon EMR version 6.12.0

Versions d’application 6.12.0

Cette version inclut les applications suivantes : Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
AWS SDK pour Java	1,12,490	1,12,446	1,12,446	1,12,397
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,15	2,1,15	2,1,15	2,1,15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.2.0	2.2.0	2.2.0
Flink	1.17.0	1.16.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4,17	2.4.15	2.4.15	2.4.15

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-0	0.13.0-amzn-0	0.13.0-amzn-0	0.12.2-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.10.0
Iceberg	1.3.0-amzn-0	1.2.0-amzn-0	1.2.0-amzn-0	1.1.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.5.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,281	0,279	0,279	0,278
Spark	3.4.0	3.3.2	3.3.2	3.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
Trino (PrestoSQL)	414	410	410	403
Zeppelin	0,1,1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notes de mise à jour 6.12.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.12.0. Les modifications ont été apportées à la version 6.11.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.12.0.

Nouvelles fonctionnalités

- Amazon EMR 6.12.0 supports Apache Spark 3.4.0, Apache Spark RAPIDS 23.06.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.1-amzn-0, Apache Iceberg 1.3.0-amzn-0, Trino 414, and PrestoDB 0.281.
- Les versions 6.12.0 et supérieures d'Amazon EMR prennent en charge l'intégration LDAP avec Apache Livy, Apache Hive via HiveServer 2 (HS2), Trino, Presto et Hue. Vous pouvez également installer Apache Spark et Apache Hadoop sur un cluster EMR utilisant la version 6.12.0 ou supérieure et les configurer pour utiliser LDAP. Pour plus d'informations, consultez Utilisation de serveurs Active Directory ou LDAP pour l'authentification avec Amazon EMR.

Modifications, améliorations et problèmes résolus

- Les versions 6.12.0 et supérieures d'Amazon EMR fournissent la prise en charge de l'exécution Java 11 pour Flink. Pour de plus amples informations, veuillez consulter Configurer Flink pour qu'il fonctionne avec Java 11.
- La version 6.12.0 ajoute un nouveau mécanisme de nouvelle tentative au flux de travail de dimensionnement des clusters pour les clusters EMR qui exécutent Presto ou Trino. Cette amélioration réduit le risque que le redimensionnement du cluster soit bloqué indéfiniment en raison de l'échec d'une seule opération de redimensionnement. Cela améliore également l'utilisation du cluster, car celui-ci augmente et diminue la capacité plus rapidement.

- La version 6.12.0 corrige un problème à cause duquel les opérations de réduction de la taille du cluster peuvent se bloquer lorsqu'un nœud principal en cours de mise hors service appropriée tombe en panne pour une raison quelconque avant d'être complètement mis hors service.
- La version 6.12.0 améliore la logique de réduction de la taille des clusters afin que votre cluster ne tente pas une réduction d'échelle des nœuds principaux en dessous du paramètre de facteur de réplication HDFS défini pour le cluster. Cela répond à vos exigences en matière de redondance des données et réduit le risque de blocage d'une opération de dimensionnement.
- La version 6.12.0 améliore les performances et l'efficacité du service de surveillance de l'état d'Amazon EMR en augmentant la vitesse à laquelle il enregistre les changements d'état des instances. Cette amélioration réduit le risque de dégradation des performances des nœuds de cluster qui exécutent plusieurs outils clients personnalisés ou des applications tierces.
- La version 6.12.0 améliore les performances du démon de gestion des journaux sur le cluster pour Amazon EMR. Par conséquent, les risques de dégradation des performances sont moindres avec les clusters EMR qui exécutent des étapes avec une simultanéité élevée.
- Avec la version 6.12.0 d'Amazon EMR, le démon de gestion des journaux a été mis à niveau pour identifier tous les journaux en cours d'utilisation avec des descripteurs de fichiers ouverts sur le stockage d'instance local, ainsi que les processus associés. Cette mise à niveau garantit qu'Amazon EMR supprime correctement les fichiers et récupère de l'espace de stockage une fois les journaux archivés dans Amazon S3.
- La version 6.12.0 inclut une amélioration du démon de gestion des journaux qui supprime les répertoires d'étapes vides et inutilisés dans le système de fichiers du cluster local. Un trop grand nombre de répertoires vides peut dégrader les performances des démons Amazon EMR et entraîner une surutilisation du disque.
- La version 6.12.0 permet la rotation des journaux pour les journaux YARN Timeline Server. Cela permet de minimiser les scénarios de surutilisation des disques, en particulier pour les clusters de longue durée.
- La taille du volume racine par défaut est passée à 15 Go dans Amazon EMR 6.10.0 et versions ultérieures. Les versions antérieures ont une taille de volume racine par défaut de 10 Go.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 822.0	4,1,4322	30 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre)

Versions Java par défaut 6.12.0

La version 6.12.0 et les versions ultérieures d'Amazon EMR prennent en charge toutes les applications dotées d'Amazon Corretto 8 par défaut, à l'exception de Trino. Pour Trino, Amazon EMR prend en charge Amazon Corretto 17 par défaut à partir de la version 6.9.0 d'Amazon EMR. Amazon EMR prend également en charge certaines applications avec Amazon Corretto 11 et 17. Ces applications sont répertoriées dans le tableau suivant. Si vous voulez modifier la JVM par défaut sur votre cluster, suivez les instructions fournies dans Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique pour chaque application exécutée sur le cluster. Vous ne pouvez utiliser qu'une seule version d'exécution de Java pour un cluster. Amazon EMR ne prend pas en charge l'exécution de différents nœuds ou applications sur différentes versions d'exécution sur le même cluster.

Bien qu'Amazon EMR prenne en charge Amazon Corretto 11 et 17 sur Apache Spark, Apache Hadoop et Apache Hive, les performances peuvent régresser pour certaines charges de travail lorsque vous utilisez ces versions de Corretto. Nous vous recommandons de tester vos charges de travail avant de modifier les valeurs par défaut.

Le tableau suivant présente les versions de Java par défaut pour les applications dans Amazon EMR 6.12.0 :

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Delta	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8

Application	Version Java/Amazon Corretto (la valeur par défaut est en gras)
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

Versions des composants 6.12.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.2	Kit de développement logiciel Amazon SageMaker Spark
delta	2.4.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
delta-standalone-connectors	0.6.0	Les connecteurs Delta fournissent différents environnements d'exécution pour intégrer Delta Lake avec des moteurs tels que Flink, Hive et Presto.
emr-ddb	5.1.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	3.5.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.9.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise

Composant	Version	Description
emr-s3-dist-cp	2.26.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.5.0	Connecteur S3Select EMR
emr-wal-cli	1.1.0	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,57,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.17.0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.17.0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	3.3.3-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.3.3-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.3-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.3.3-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.3.3-amzn-4	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.3-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.3-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.3-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	3.3.3-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.3.3-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.17-amzn-0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.17-amzn-0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.17-amzn-0	HBase client en ligne de commande.
hbase-rest-server	2.4.17-amzn-0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.17-amzn-0	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.17-amzn-0	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-5	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	3.1.3-amzn-5	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-5	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-5	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-5	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-5	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-5	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.13.1-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.13.1-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.13.1-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.

Composant	Version	Description
hudi-spark	0.13.1-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1.3.0-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,8.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.7.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.3	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	5.1.3	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.281-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.281-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.281-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	414-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.

Composant	Version	Description
trino-worker	414-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	414-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.4.0-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.4.0-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.4.0-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.4.0-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	23.06.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.

Composant	Version	Description
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-3	Les bibliothèques et l'applica tion tez YARN.
tez-on-worker	0.10.2-amzn-3	Les bibliothèques et l'applica tion tez YARN pour les nœuds de travail.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 6.12.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.12.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Modification de la page d'accueil Java KMS d'Hadoop	Not available.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Classifications	Description	Actions de reconfiguration
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Classifications	Description	Actions de reconfiguration
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.

Classifications	Description	Actions de reconfiguration
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

Journal des modifications 6.12.0

Journal des modifications pour la version 6.12.0 et les notes de mise à jour

Date	Événement	Description
27/07/2023	Documentation mise à jour	Mise à jour des options Java pour la version 6.12 et ajout du tutoriel Oozie pour la mise à jour de la JVM
21/07/2023	Déploiement terminé	Amazon EMR 6.12.0 entièrement déployé dans toutes les régions prises en charge
21/07/2023	Publication de documents	Notes de mise à jour d'Amazon EMR 6.12.0 publiées pour la première fois
12/07/2023	Première version	Amazon EMR 6.12.0 a été déployé pour la première fois dans des régions commercia les limitées

Amazon EMR version 6.11.1

Versions d’application 6.11.1

Cette version inclut les applications suivantes : Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
AWS SDK pour Java	1,12,446	1,12,446	1,12,397	1,12,397
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,15	2,1,15	2,1,15	2,1,15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.2.0	2.2.0
Flink	1.16.0	1.16.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.15	2.4.15

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.0-amzn-0	0.13.0-amzn-0	0.12.2-amzn-0	0.12.2-amzn-0
Hue	4.11.0	4.11.0	4.10.0	4.10.0
Iceberg	1.2.0-amzn-0	1.2.0-amzn-0	1.1.0-amzn-0	1.1.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.4.1	1.4.1	1.5.0	1.5.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,279	0,279	0,278	0,278
Spark	3.3.2	3.3.2	3.3.1	3.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
Trino (PrestoSQL)	410	410	403	403
Zeppelin	0,1,1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notes de mise à jour 6.11.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.11.1. Les modifications ont été apportées à la version 6.11.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.11.1.

Modifications, améliorations et problèmes résolus

- En raison d'un conflit de verrouillage, un nœud peut se retrouver bloqué s'il est ajouté ou supprimé en même temps qu'il est mis hors service. Par conséquent, le gestionnaire de ressources Hadoop (YARN) ne répond plus et affecte tous les conteneurs entrants et en cours d'exécution.
- Cette version inclut une modification qui permet aux clusters à haute disponibilité de se remettre d'un état défaillant après le redémarrage.
- Cette version inclut des correctifs de sécurité pour Hue et HBase.
- Cette version résout un problème selon lequel les clusters exécutant des charges de travail sur Spark avec Amazon EMR peuvent recevoir silencieusement des résultats incorrects avec `contains`, `startsWith`, `endsWith` et `like`. Ce problème se produit lorsque vous utilisez les expressions sur des champs partitionnés contenant des métadonnées dans Amazon EMR Hive3 Metastore Server (HMS).
- Cette version corrige un problème de limitation du côté de Glue en l'absence de fonctions définies par l'utilisateur (UDF).
- Cette version corrige un problème qui supprime les journaux des conteneurs par le service d'agrégation des journaux des nœuds avant que le transmetteur de journaux ne puisse les envoyer vers S3 en cas de mise hors service de YARN.
- Cette version corrige un problème lié aux métriques du FairShare planificateur lorsque Node Label est activé pour Hadoop.

- Cette version corrige un problème qui affectait les performances de Spark lorsque vous définissez une valeur `true` par défaut pour la configuration `spark.yarn.heterogeneousExecutors.enabled` dans `spark-defaults.conf`.
- Cette version corrige un problème lié à l'échec de la lecture des données de shuffle par Reduce Task. Ce problème provoquait des échecs de requêtes Hive avec une erreur de mémoire corrompue.
- Cette version ajoute un nouveau mécanisme de nouvelle tentative au flux de travail de dimensionnement des clusters pour les clusters EMR qui exécutent Presto ou Trino. Cette amélioration réduit le risque que le redimensionnement du cluster soit bloqué indéfiniment en raison de l'échec d'une seule opération de redimensionnement. Cela améliore également l'utilisation du cluster, car celui-ci augmente et diminue la capacité plus rapidement.
- Cette version améliore la logique de réduction de la taille des clusters afin que votre cluster ne tente pas une réduction d'échelle des nœuds principaux en dessous du paramètre de facteur de réplication HDFS défini pour le cluster. Cela répond à vos exigences en matière de redondance des données et réduit le risque de blocage d'une opération de dimensionnement.
- Le démon de gestion des journaux a été mis à niveau pour identifier tous les journaux en cours d'utilisation avec des descripteurs de fichiers ouverts sur le stockage d'instance local, ainsi que les processus associés. Cette mise à niveau garantit qu'Amazon EMR supprime correctement les fichiers et récupère de l'espace de stockage une fois les journaux archivés dans Amazon S3.
- Cette version inclut une amélioration du démon de gestion des journaux qui supprime les répertoires d'étapes vides et inutilisés dans le système de fichiers du cluster local. Un trop grand nombre de répertoires vides peut dégrader les performances des démons Amazon EMR et entraîner une surutilisation du disque.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique (Taipei), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Jakarta) (Malaisie), Asie-Pacifique (Melbourne), Asie-Paci fique (Mumbai), Asie-Paci fique (Osaka), Asie-Paci fique (Séoul), Asie-Paci fique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Malaisie), Asie-Pacifique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Émirats arabes unis), Canada (centre),

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Canada Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Calgary), Europe (Zurich), Asie-Pacifique (Malaisie)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 822.0	4,1,4322	30 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

Versions des composants 6.11.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon

EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.2	Kit de développement logiciel Amazon SageMaker Spark
delta	2.2.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
delta-standalone-connectors	0.6.0	Les connecteurs Delta fournissent différents environnements d'exécution pour intégrer Delta Lake avec des moteurs tels que Flink, Hive et Presto.
emr-ddb	5.1.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	3.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.8.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.25.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.4.0	Connecteur S3Select EMR
emr-wal-cli	1.1.0	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,56,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.16.0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.16.0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.3.3-amzn-3.1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.3.3-amzn-3.1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.3-amzn-3.1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.3.3-amzn-3.1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.3.3-amzn-3.1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.3-amzn-3.1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.3-amzn-3.1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-3.1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	3.3.3-amzn-3.1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.3.3-amzn-3.1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.3.3-amzn-3.1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.15-amzn-1.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.15-amzn-1.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.15-amzn-1.1	HBase client en ligne de commande.
hbase-rest-server	2.4.15-amzn-1.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.15-amzn-1.1	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.15-amzn-1.1	Outil de réparation pour les HBase clusters Apache.

Composant	Version	Description
hcatalog-client	3.1.3-amzn-4.1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-4.1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-4.1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-4.1	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-4.1	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-4.1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-4.1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.13.0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.13.0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.

Composant	Version	Description
hudi-trino	0.13.0-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.13.0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1.2.0-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,8.0	Pilotes Nvidia et boîte à outils Cuda

Composant	Version	Description
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.2	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	5.1.2	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.279-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.279-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.279-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.

Composant	Version	Description
trino-coordinator	410-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	410-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	410-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.3.2-amzn-0.1	Clients de ligne de commande Spark.
spark-history-server	3.3.2-amzn-0.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.3.2-amzn-0.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.3.2-amzn-0.1	Bibliothèques Apache Spark requises par les esclaves YARN.

Composant	Version	Description
spark-rapids	23.02.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-2.1	Les bibliothèques et l'applica tion tez YARN.
tez-on-worker	0.10.2-amzn-2.1	Les bibliothèques et l'applica tion tez YARN pour les nœuds de travail.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 6.11.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.11.1

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop

Classifications	Description	Actions de reconfiguration
		Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Classifications	Description	Actions de reconfiguration
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Classifications	Description	Actions de reconfiguration
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.

Classifications	Description	Actions de reconfiguration
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

Journal des modifications 6.11.1

Journal des modifications pour la version 6.11.1 et les notes de mise à jour

Date	Événement	Description
30/08/2023	Mise à jour des notes de mise à jour	Ajout de plusieurs corrections liées au plan de contrôle dans les notes de mise à jour
21/08/2023	Publication de documents	Notes de mise à jour d'Amazon EMR 6.11.1 publiées pour la première fois
16/08/2023	Déploiement terminé	Amazon EMR 6.11.1 entièrement déployé dans toutes les régions prises en charge
04/08/2023	Première version	Amazon EMR 6.11.1 a été déployé pour la première fois dans des régions commercia les limitées

Amazon EMR version 6.11.0

Versions d’application 6.11.0

Cette version inclut les applications suivantes : Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.11.0	emr-6.10.1	emr-6.10.0	emr-6.9.1
AWS SDK pour Java	1,12,446	1,12,397	1,12,397	1,12,170
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,15	2,1,15	2,1,15	2,1,15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.2.0	2.1.0
Flink	1.16.0	1.16.0	1.16.0	1.15.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.15	2.4.13

	emr-6.11.0	emr-6.10.1	emr-6.10.0	emr-6.9.1
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.0-amzn-0	0.12.2-amzn-0	0.12.2-amzn-0	0.12.1-amzn-0
Hue	4.11.0	4.10.0	4.10.0	4.10.0
Iceberg	1.2.0-amzn-0	1.1.0-amzn-0	1.1.0-amzn-0	0.14.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.4.1	1.5.0	1.5.0	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,279	0,278	0,278	0,276
Spark	3.3.2	3.3.1	3.3.1	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.10.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2

	emr-6.11.0	emr-6.10.1	emr-6.10.0	emr-6.9.1
Trino (PrestoSQL)	410	403	403	398
Zeppelin	0,1,1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notes de mise à jour 6.11.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.11.0. Les modifications ont été apportées à la version 6.10.0. Pour plus d'informations sur le calendrier de publication, consultez le journal des modifications.

Nouvelles fonctionnalités

- Amazon EMR 6.11.0 prend en charge Apache Spark 3.3.2-amzn-0, Apache Spark RAPIDS 23.02.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.0-amzn-0, Apache Iceberg 1.2.0-amzn-0, Trino 410-amzn-0 et PrestoDB 0.279-amzn-0.

Modifications, améliorations et problèmes résolus

- Avec Amazon EMR 6.11.0, le connecteur DynamoDB a été mis à niveau vers la version 5.0.0. La version 5.0.0 utilise AWS SDK for Java 2.x. Les versions précédentes utilisaient la version AWS SDK pour Java 1.x. À la suite de cette mise à niveau, nous vous conseillons vivement de tester votre code avant d'utiliser le connecteur DynamoDB avec Amazon EMR 6.11.
- Lorsque le connecteur DynamoDB pour Amazon EMR 6.11.0 appelle le service DynamoDB, il utilise la valeur de région que vous fournissez pour la propriété `dynamodb.endpoint`. Nous vous recommandons de configurer également `dynamodb.region` lorsque vous utilisez `dynamodb.endpoint`, et que les deux propriétés ciblent la même Région AWS. Si vous l'utilisez `dynamodb.endpoint` et que vous ne le configurez pas`dynamodb.region`, le connecteur DynamoDB pour Amazon EMR 6.11.0 renverra une exception de région non valide et tentera de réconcilier vos Région AWS informations avec le service de métadonnées d'instance Amazon EC2 (IMDS). Si le connecteur ne parvient pas à récupérer la région depuis IMDS, la valeur par défaut est USA Est (Virginie du Nord) (`us-east-1`). L'erreur suivante illustre l'exception de région non valide que vous pourriez

obtenir si vous ne configurez pas correctement la `dynamodb.region` propriété : `error software.amazon.awssdk.services.dynamodb.model.DynamoDbException: Credential should be scoped to a valid region.` pour plus d'informations sur les classes concernées par la AWS SDK pour Java mise à niveau vers la version 2.x, consultez le commit Upgrade AWS SDK pour Java from 1.x to 2.x (#175) dans le GitHub dépôt du connecteur Amazon EMR - DynamoDB.

- Cette version corrige un problème où les données des colonnes deviennent `NULL` lorsque vous utilisez Delta Lake pour stocker les données de la table Delta dans Amazon S3 après l'opération de changement de nom de colonne. Pour plus d'informations sur cette fonctionnalité expérimentale dans Delta Lake, consultez Opération de changement de nom de colonne dans le guide de l'utilisateur Delta Lake.
- La version 6.11.0 résout un problème qui peut survenir lorsque vous créez un nœud périphérique en répliquant l'un des nœuds primaires à partir d'un cluster comportant plusieurs nœuds primaires. Le nœud périphérique répliqué peut retarder les opérations de réduction d'échelle ou entraîner une utilisation élevée de la mémoire sur les nœuds primaires. Pour plus d'informations sur la création d'un nœud périphérique pour communiquer avec votre cluster EMR, consultez la section Edge Node Creator dans le `aws-samples` référentiel sur. GitHub
- La version 6.11.0 améliore le processus d'automatisation utilisé par Amazon EMR pour remonter les volumes Amazon EBS sur une instance après un redémarrage.
- La version 6.11.0 corrige un problème qui entraînait des écarts intermittents dans les métriques Hadoop publiées par Amazon EMR sur Amazon. CloudWatch
- La version 6.11.0 résout un problème lié aux clusters EMR où une mise à jour du fichier de configuration YARN contenant la liste d'exclusion des nœuds du cluster est interrompue en raison d'une surutilisation du disque. La mise à jour incomplète entrave les futures opérations de réduction de la taille du cluster. Cette version garantit que votre cluster reste sain et que les opérations de dimensionnement fonctionnent comme prévu.
- La taille du volume racine par défaut est passée à 15 Go dans Amazon EMR 6.10.0 et versions ultérieures. Les versions antérieures ont une taille de volume racine par défaut de 10 Go.
- Hadoop 3.3.3 a introduit une modification dans YARN (YARN-9608) qui maintient les nœuds sur lesquels les conteneurs s'exécutaient dans un état de mise hors service jusqu'à ce que l'application soit terminée. Cette modification permet de s'assurer que les données locales telles que les données réorganisées ne sont pas perdues et que vous n'avez pas besoin de réexécuter la tâche. Cette approche peut également entraîner une sous-utilisation des ressources sur les clusters avec ou sans activation de la mise à l'échelle gérée.

Dans les versions 6.11.0 et supérieures d'Amazon EMR ainsi que dans les versions 6.8.1, 6.9.1 et 6.10.1, la valeur de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` est définie sur `false` dans `yarn-site.xml` pour résoudre ce problème.

Bien que ce correctif règle les problèmes introduits par YARN-9608, il peut entraîner l'échec des tâches Hive en raison de la perte de données de réorganisation sur les clusters pour lesquels la mise à l'échelle gérée est activée. Nous avons atténué ce risque dans cette version en définissant également `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` pour les charges de travail Hive. Cette configuration n'est disponible qu'à partir de la version 6.11.0 d'Amazon EMR.

- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

> **Note**
>
> Cette version ne bénéficie plus de mises à jour automatiques de l'AMI puisqu'elle a été suivie d'une version supplémentaire de correctifs. La version du correctif est indiquée par le numéro qui suit la deuxième décimale (`6.8.1`). Pour savoir si vous utilisez la dernière version du correctif, consultez les versions disponibles dans le Guide des versions, ou consultez le menu déroulant des versions d'Amazon EMR lorsque vous créez un cluster dans la console, ou utilisez l'action d'API `ListReleaseLabels` ou de CLI `list-release-labels`. Pour être tenu au courant des nouvelles versions, abonnez-vous au flux RSS sur la page Quoi de neuf ?

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Zurich), Europe (Milan), Europe (Espagne), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Zurich), Europe (Milan), Europe (Espagne), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre)

Versions des composants 6.11.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.2	Kit de développement logiciel Amazon SageMaker Spark
delta	2.2.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
delta-standalone-connectors	0.6.0	Les connecteurs Delta fournissent différents environnements d'exécution pour intégrer Delta Lake avec des moteurs tels que Flink, Hive et Presto.
emr-ddb	5.1.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.

Composant	Version	Description
emr-goodies	3.4.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.8.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.25.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.4.0	Connecteur S3Select EMR
emr-wal-cli	1.1.0	CLI utilisée pour la liste/sup pression EMR WAL.
emrfs	2,56,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.16.0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.16.0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.3.3-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.3.3-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.3-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.3.3-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.3.3-amzn-3	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.3-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.3-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	3.3.3-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.3-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.3.3-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.3.3-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.15-amzn-1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.15-amzn-1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.15-amzn-1	HBase client en ligne de commande.
hbase-rest-server	2.4.15-amzn-1	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	2.4.15-amzn-1	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.15-amzn-1	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-4	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-4	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-4	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-4	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-4	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-4	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-4	Service pour l'acceptation de requêtes Hive en tant que requêtes web.

Composant	Version	Description
hudi	0.13.0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.13.0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.13.0-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.13.0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	1.2.0-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse

Composant	Version	Description
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,8.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.2	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	5.1.2	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.279-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.279-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.279-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	410-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	410-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	410-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.3.2-amzn-0	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	3.3.2-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.3.2-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.3.2-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	23.02.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-2	Les bibliothèques et l'applica tion tez YARN.
tez-on-worker	0.10.2-amzn-2	Les bibliothèques et l'applica tion tez YARN pour les nœuds de travail.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 6.11.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.11.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classifications	Description	Actions de reconfiguration
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.

Classifications	Description	Actions de reconfiguration
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.pr operties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-s ite.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.pr operties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.prope rties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.

Classifications	Description	Actions de reconfiguration
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.

Classifications	Description	Actions de reconfiguration
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.p roperties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.

Classifications	Description	Actions de reconfiguration
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Journal des modifications 6.11.0

Journal des modifications pour la version 6.11.0 et les notes de mise à jour

Date	Événement	Description
21/08/2023	Mettre à jour	Correction d'un problème introduit avec Hadoop 3.3.3.
26/07/2023	Mettre à jour	Étiquettes `2.0.20230612.0` et `2.0.20230628.0` de la nouvelle version du système d'exploitation.
09/06/2023	Déploiement terminé	Amazon EMR 6.11.0 entièrement déployé dans toutes les régions prises en charge
09/06/2023	Publication de documents	Notes de mise à jour d'Amazon EMR 6.11.0 publiées pour la première fois
08/06/2023	Première version	Amazon EMR 6.11.0 a été déployé pour la première fois dans des régions commercia les limitées

Amazon EMR version 6.10.1

Versions d’application 6.10.1

Cette version inclut les applications suivantes : Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.10.1	emr-6.10.0	emr-6.9.1	emr-6.9.0
AWS SDK pour Java	1,12,397	1,12,397	1,12,170	1,12,170
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,15	2,1,15	2,1,15	2,1,15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.1.0	2.1.0
Flink	1.16.0	1.16.0	1.15.2	1.15.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.13	2.4.13
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.2-amzn-0	0.12.2-amzn-0	0.12.1-amzn-0	0.12.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0

	emr-6.10.1	emr-6.10.0	emr-6.9.1	emr-6.9.0
Iceberg	1.1.0-amzn-0	1.1.0-amzn-0	0.14.1-amzn-0	0.14.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,278	0,278	0,276	0,276
Spark	3.3.1	3.3.1	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.10.0	2.10.0
Tez	0,1,2	0,1,2	0,1,2	0,1,2
Trino (PrestoSQL)	403	403	398	398
Zeppelin	0,1,1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notes de mise à jour 6.10.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.10.1. Les modifications ont été apportées à la version 6.10.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.10.1.

Modifications, améliorations et problèmes résolus

- En raison d'un conflit de verrouillage, un nœud peut se retrouver bloqué s'il est ajouté ou supprimé en même temps qu'il est mis hors service. Par conséquent, le gestionnaire de ressources Hadoop (YARN) ne répond plus et affecte tous les conteneurs entrants et en cours d'exécution.
- Hadoop 3.3.3 a introduit une modification dans YARN (YARN-9608) qui maintient les nœuds sur lesquels les conteneurs s'exécutaient dans un état de mise hors service jusqu'à ce que l'application soit terminée. Cette modification permet de s'assurer que les données locales telles que les données réorganisées ne sont pas perdues et que vous n'avez pas besoin de réexécuter la tâche. Cette approche peut également entraîner une sous-utilisation des ressources sur les clusters avec ou sans activation de la mise à l'échelle gérée.

  Dans les versions 6.11.0 et supérieures d'Amazon EMR ainsi que dans les versions 6.8.1, 6.9.1 et 6.10.1, la valeur de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` est définie sur `false` dans `yarn-site.xml` pour résoudre ce problème.

  Bien que ce correctif règle les problèmes introduits par YARN-9608, il peut entraîner l'échec des tâches Hive en raison de la perte de données de réorganisation sur les clusters pour lesquels la mise à l'échelle gérée est activée. Nous avons atténué ce risque dans cette version en définissant également `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` pour les charges de travail Hive. Cette configuration n'est disponible qu'à partir de la version 6.11.0 d'Amazon EMR.
- Le collecteur de métriques n'enverra aucune métrique au plan de contrôle après le basculement du nœud primaire dans les clusters avec la configuration des groupes d'instances.
- Cette version inclut une modification qui permet aux clusters à haute disponibilité de se remettre d'un état défaillant après le redémarrage.
- Cette version inclut des correctifs de sécurité pour Hue et HBase.
- Cette version résout un problème selon lequel les clusters exécutant des charges de travail sur Spark avec Amazon EMR peuvent recevoir silencieusement des résultats incorrects avec `contains`, `startsWith`, `endsWith` et `like`. Ce problème se produit lorsque vous utilisez les

expressions sur des champs partitionnés contenant des métadonnées dans Amazon EMR Hive3 Metastore Server (HMS).

- Cette version corrige un problème de limitation du côté de Glue en l'absence de fonctions définies par l'utilisateur (UDF).
- Cette version corrige un problème qui supprime les journaux des conteneurs par le service d'agrégation des journaux des nœuds avant que le transmetteur de journaux ne puisse les envoyer vers S3 en cas de mise hors service de YARN.
- Cette version corrige un problème lié aux métriques du FairShare planificateur lorsque Node Label est activé pour Hadoop.
- Cette version corrige un problème qui affectait les performances de Spark lorsque vous définissez une valeur `true` par défaut pour la configuration `spark.yarn.heterogeneousExecutors.enabled` dans `spark-defaults.conf`.
- Cette version corrige un problème lié à l'échec de la lecture des données de shuffle par Reduce Task. Ce problème provoquait des échecs de requêtes Hive avec une erreur de mémoire corrompue.
- Cette version ajoute un nouveau mécanisme de nouvelle tentative au flux de travail de dimensionnement des clusters pour les clusters EMR qui exécutent Presto ou Trino. Cette amélioration réduit le risque que le redimensionnement du cluster soit bloqué indéfiniment en raison de l'échec d'une seule opération de redimensionnement. Cela améliore également l'utilisation du cluster, car celui-ci augmente et diminue la capacité plus rapidement.
- Cette version améliore la logique de réduction de la taille des clusters afin que votre cluster ne tente pas une réduction d'échelle des nœuds principaux en dessous du paramètre de facteur de réplication HDFS défini pour le cluster. Cela répond à vos exigences en matière de redondance des données et réduit le risque de blocage d'une opération de dimensionnement.
- Le démon de gestion des journaux a été mis à niveau pour identifier tous les journaux en cours d'utilisation avec des descripteurs de fichiers ouverts sur le stockage d'instance local, ainsi que les processus associés. Cette mise à niveau garantit qu'Amazon EMR supprime correctement les fichiers et récupère de l'espace de stockage une fois les journaux archivés dans Amazon S3.
- Cette version inclut une amélioration du démon de gestion des journaux qui supprime les répertoires d'étapes vides et inutilisés dans le système de fichiers du cluster local. Un trop grand nombre de répertoires vides peut dégrader les performances des démons Amazon EMR et entraîner une surutilisation du disque.
- Cette version résout un problème qui peut survenir lorsque vous créez un nœud périphérique en répliquant l'un des nœuds primaires à partir d'un cluster comportant plusieurs nœuds primaires. Le

nœud périphérique répliqué peut retarder les opérations de réduction d'échelle ou entraîner une utilisation élevée de la mémoire sur les nœuds primaires. Pour plus d'informations sur la création d'un nœud périphérique pour communiquer avec votre cluster EMR, consultez la section Edge Node Creator dans le `aws-samples` référentiel sur. GitHub

- Cette version améliore le processus d'automatisation utilisé par Amazon EMR pour remonter les volumes Amazon EBS sur une instance après un redémarrage.
- Cette version corrige un problème qui entraînait des écarts intermittents dans les métriques Hadoop publiées par Amazon EMR sur Amazon. CloudWatch
- Cette version résout un problème lié aux clusters EMR où une mise à jour du fichier de configuration YARN contenant la liste d'exclusion des nœuds du cluster est interrompue en raison d'une surutilisation du disque. La mise à jour incomplète entrave les futures opérations de réduction de la taille du cluster. Cette version garantit que votre cluster reste sain et que les opérations de dimensionnement fonctionnent comme prévu.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 822.0	4,1,4322	30 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

Versions des composants 6.10.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon

EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.2	Kit de développement logiciel Amazon SageMaker Spark
delta	2.2.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.3.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.7.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.24.0	Application de copie distribuée optimisée pour Amazon S3.

Composant	Version	Description
emr-s3-select	2.3.0	Connecteur S3Select EMR
emr-wal-cli	1.0.0	CLI utilisée pour la liste/suppression EMR WAL.
emrfs	2,55,0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.16.0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.16.0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.3.3-amzn-2.1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».

Composant	Version	Description
hadoop-hdfs-datanode	3.3.3-amzn-2.1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.3-amzn-2.1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.3.3-amzn-2.1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.3.3-amzn-2.1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.3-amzn-2.1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.3-amzn-2.1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-2.1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.3-amzn-2.1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.3.3-amzn-2.1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	3.3.3-amzn-2.1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.15-amzn-0.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.15-amzn-0.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.15-amzn-0.1	HBase client en ligne de commande.
hbase-rest-server	2.4.15-amzn-0.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.15-amzn-0.1	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.15-amzn-0.1	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-3.1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-3.1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.

Composant	Version	Description
hcatalog-webhcat-server	3.1.3-amzn-3.1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-3.1	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-3.1	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-3.1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-3.1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.12.2-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.12.2-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.12.2-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.12.2-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.

Composant	Version	Description
hue-server	4.10.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
iceberg	1.1.0-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,8.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur

Composant	Version	Description
phoenix-library	5.1.2	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	6.0.0-SNAPSHOT	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.278.1-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.278.1-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.278.1-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	403-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	403-amzn-0	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
trino-client	403-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.3.1-amzn-0.1	Clients de ligne de commande Spark.
spark-history-server	3.3.1-amzn-0.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.3.1-amzn-0.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.3.1-amzn-0.1	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	22.12.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	2.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-1.1	Les bibliothèques et l'application tez YARN.
tez-on-worker	0.10.2-amzn-1.1	Les bibliothèques et l'application tez YARN pour les nœuds de travail.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 6.10.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions

de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.10.1

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode,

Classifications	Description	Actions de reconfiguration
		SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Classifications	Description	Actions de reconfiguration
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Classifications	Description	Actions de reconfiguration
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.

Classifications	Description	Actions de reconfiguration
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

Journal des modifications 6.10.1

Journal des modifications pour la version 6.10.1 et les notes de mise à jour

Date	Événement	Description
30/08/2023	Mise à jour des notes de mise à jour	Ajout de plusieurs corrections liées au plan de contrôle dans les notes de mise à jour
21/08/2023	Publication de documents	Notes de mise à jour d'Amazon EMR 6.10.1 publiées pour la première fois
16/08/2023	Déploiement terminé	Amazon EMR 6.10.1 entièrement déployé dans toutes les régions prises en charge
04/08/2023	Première version	Amazon EMR 6.10.1 a été déployé pour la première fois dans des régions commercia les limitées

Amazon EMR version 6.10.0

Versions d’application 6.10.0

Cette version inclut les applications suivantes : Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.10.0	emr-6.9.1	emr-6.9.0	emr-6.8.1
AWS SDK pour Java	1,12,397	1,12,170	1,12,170	1,12,170
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,15	2,1,15	2,1,15	2,1,15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.1.0	2.1.0	-
Flink	1.16.0	1.15.2	1.15.2	1.15.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.13	2.4.13	2.4.12

	emr-6.10.0	emr-6.9.1	emr-6.9.0	emr-6.8.1
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.2-amzn-0	0.12.1-amzn-0	0.12.1-amzn-0	0.11.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	1.1.0-amzn-0	0.14.1-amzn-0	0.14.1-amzn-0	0.14.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.1.0
JupyterHub	1.5.0	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,278	0,276	0,276	0,273
Spark	3.3.1	3.3.0	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.10.0	2.10.0	2.9.1
Tez	0,1,2	0,1,2	0,1,2	0.9.2

	emr-6.10.0	emr-6.9.1	emr-6.9.0	emr-6.8.1
Trino (PrestoSQL)	403	398	398	388
Zeppelin	0,1,1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notes de mise à jour 6.10.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.10.0. Les modifications ont été apportées à la version 6.9.0. Pour plus d'informations sur le calendrier de publication, consultez le journal des modifications.

Nouvelles fonctionnalités

- Amazon EMR 6.10.0 prend en charge Apache Spark 3.3.1, Apache Spark RAPIDS 22.12.0, CUDA 11.8.0, Apache Hudi 0.12.2-amzn-0, Apache Iceberg 1.1.0-amzn-0, Trino 403 et PrestoDB 0.278.1.
- Amazon EMR 6.10.0 inclut un connecteur Trino-Hudi natif qui fournit un accès en lecture aux données des tables Hudi. Vous pouvez activer le connecteur avec `trino-cli --catalog hudi` et le configurer en fonction de vos besoins avec `trino-connector-hudi`. L'intégration native avec Amazon EMR signifie que vous n'avez plus besoin d'utiliser `trino-connector-hive` pour interroger les tables Hudi. Pour obtenir la liste des configurations prises en charge avec le nouveau connecteur, consultez la page Connecteur Hudi dans la documentation de Trino.
- Les versions 6.10.0 et supérieures d'Amazon EMR prennent en charge l'intégration d'Apache Zeppelin avec Apache Flink. Pour plus d’informations, consultez Travailler avec les jobs Flink de Zeppelin dans Amazon EMR.

Problèmes connus

- Hadoop 3.3.3 a introduit une modification dans YARN (YARN-9608) qui maintient les nœuds sur lesquels les conteneurs s'exécutaient dans un état de mise hors service jusqu'à ce que l'application soit terminée. Cette modification permet de s'assurer que les données locales telles que les données réorganisées ne sont pas perdues et que vous n'avez pas besoin de réexécuter la tâche.

Cette approche peut également entraîner une sous-utilisation des ressources sur les clusters avec ou sans activation de la mise à l'échelle gérée.

Pour contourner ce problème dans Amazon EMR 6.10.0, vous pouvez définir la valeur de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` sur `false` dans `yarn-site.xml`. Dans les versions 6.11.0 et supérieures d'Amazon EMR, ainsi que 6.8.1, 6.9.1 et 6.10.1, la configuration est définie sur `false` par défaut pour résoudre ce problème.

Modifications, améliorations et problèmes résolus

- Amazon EMR 6.10.0 supprime la dépendance sur `minimal-json.jar` pour l'intégration Amazon Redshift pour Apache Spark et ajoute automatiquement les fichiers JAR liés à Spark-Redshift requis au chemin de classe de l'exécuteur pour Spark : `spark-redshift.jar`, `spark-avro.jar` et `RedshiftJDBC.jar`.
- La version 6.10.0 améliore le démon de gestion des journaux sur le cluster afin de surveiller des dossiers de journaux supplémentaires dans votre cluster EMR. Cette amélioration permet de minimiser les scénarios de surutilisation des disques.
- La version 6.10.0 redémarre automatiquement le démon de gestion des journaux du cluster lorsqu'il s'arrête. Cette amélioration réduit le risque que les nœuds apparaissent défectueux en raison d'une surutilisation du disque.
- Amazon EMR 6.10.0 prend en charge les points de terminaison régionaux pour le mappage des utilisateurs EMRFS.
- La taille du volume racine par défaut est passée à 15 Go dans Amazon EMR 6.10.0 et versions ultérieures. Les versions antérieures ont une taille de volume racine par défaut de 10 Go.
- La version 6.10.0 corrige un problème qui provoquait le blocage des tâches Spark lorsque tous les exécuteurs Spark restants se trouvaient sur un hôte en cours de mise hors service avec le gestionnaire de ressources YARN.
- Avec Amazon EMR 6.6.0 à 6.9.x, les requêtes INSERT avec partition dynamique et clause ORDER BY ou SORT BY auront toujours deux réducteurs. Ce problème est dû à la modification d'OSS HIVE-20703, qui place l'optimisation des partitions dynamiques de tri dans le cadre d'une décision basée sur les coûts. Si votre charge de travail ne nécessite pas le tri des partitions dynamiques, nous vous recommandons de définir la propriété `hive.optimize.sort.dynamic.partition.threshold` sur `-1` pour désactiver la nouvelle fonctionnalité et obtenir le nombre de réducteurs correctement calculé. Ce problème est résolu dans OSS Hive dans le cadre de HIVE-22269 et dans Amazon EMR 6.10.0.

- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

> **Note**
>
> Cette version ne bénéficie plus de mises à jour automatiques de l'AMI puisqu'elle a été suivie d'une version supplémentaire de correctifs. La version du correctif est indiquée par le numéro qui suit la deuxième décimale (`6.8.1`). Pour savoir si vous utilisez la dernière version du correctif, consultez les versions disponibles dans le Guide des versions, ou consultez le menu déroulant des versions d'Amazon EMR lorsque vous créez un cluster dans la console, ou utilisez l'action d'API `ListReleaseLabels` ou de CLI `list-release-labels`. Pour être tenu au courant des nouvelles versions, abonnez-vous au flux RSS sur la page Quoi de neuf ?

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Zurich), Europe (Milan), Europe (Espagne), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Zurich), Europe (Milan), Europe (Espagne), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Zurich), Europe (Milan), Europe (Espagne), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 418,0	4,14.311	3 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Zurich), Europe (Milan), Europe (Espagne), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 404.1	4,14.311	18 avril 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)
2,0.2023 404.0	4,14.311	10 avril 2023	USA Est (Virginie du Nord), Europe (Paris)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 320,0	4,14,309	30 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 207.0	4,14,304	22 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

Versions des composants 6.10.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon

EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.2	Kit de développement logiciel Amazon SageMaker Spark
delta	2.2.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.3.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.7.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.24.0	Application de copie distribuée optimisée pour Amazon S3.

Composant	Version	Description
emr-s3-select	2.3.0	Connecteur S3Select EMR
emr-wal-cli	1.0.0	CLI utilisée pour la liste/suppression EMR WAL.
emrfs	2,55,0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.16.0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.16.0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.3.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».

Composant	Version	Description
hadoop-hdfs-datanode	3.3.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.3.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.3.3-amzn-2	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.3.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	3.3.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.15-amzn-0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.15-amzn-0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.15-amzn-0	HBase client en ligne de commande.
hbase-rest-server	2.4.15-amzn-0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.15-amzn-0	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.15-amzn-0	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-3	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-3	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.

Composant	Version	Description
hcatalog-webhcat-server	3.1.3-amzn-3	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-3	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-3	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-3	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-3	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.12.2-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.12.2-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.12.2-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.12.2-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.

Composant	Version	Description
hue-server	4.10.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
iceberg	1.1.0-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.5.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,8.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur

Composant	Version	Description
phoenix-library	5.1.2	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	6.0.0-SNAPSHOT	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.278.1-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.278.1-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.278.1-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	403-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	403-amzn-0	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
trino-client	403-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.3.1-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.3.1-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.3.1-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.3.1-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	22.12.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	2.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.10.2-amzn-1	Les bibliothèques et l'application tez YARN.
tez-on-worker	0.10.2-amzn-1	Les bibliothèques et l'application tez YARN pour les nœuds de travail.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 6.10.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions

de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.10.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode,

Classifications	Description	Actions de reconfiguration
		SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Classifications	Description	Actions de reconfiguration
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Modifiez les valeurs dans le fichier hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Modifiez les valeurs dans le fichier kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Classifications	Description	Actions de reconfiguration
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.

Classifications	Description	Actions de reconfiguration
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

Journal des modifications 6.10.0

Journal des modifications pour la version 6.10.0 et les notes de mise à jour

Date	Événement	Description
21/08/2023	Mettre à jour	Ajout d'un problème connu introduit avec Hadoop 3.3.3.
26/07/2023	Mettre à jour	Étiquettes `2.0.20230612.0` et `2.0.20230628.0` de la nouvelle version du système d'exploitation.
02	Déploiement terminé	Amazon EMR 6.10 entièrement déployé dans toutes les régions prises en charge
02	Publication de documents	Notes de mise à jour d'Amazon EMR 6.10 publiées pour la première fois
27/02/2023	Première version	Amazon EMR 6.10 déployé dans des régions commercia les limitées

Amazon EMR version 6.9.1

Versions d’application 6.9.1

Cette version inclut les applications suivantes : Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.9.1	emr-6.9.0	emr-6.8.1	emr-6.8.0
AWS SDK pour Java	1,12,170	1,12,170	1,12,170	1,12,170
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,15	2,1,15	2,1,15	2,1,15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.1.0	2.1.0	-	-
Flink	1.15.2	1.15.2	1.15.1	1.15.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.13	2.4.13	2.4.12	2.4.12

	emr-6.9.1	emr-6.9.0	emr-6.8.1	emr-6.8.0
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.1-amzn-0	0.12.1-amzn-0	0.11.1-amzn-0	0.11.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.1-amzn-0	0.14.1-amzn-0	0.14.0-amzn-0	0.14.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,276	0,276	0,273	0,273
Spark	3.3.0	3.3.0	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.10.0	2.10.0	2.9.1	2.9.1
Tez	0,1,2	0,1,2	0.9.2	0.9.2

	emr-6.9.1	emr-6.9.0	emr-6.8.1	emr-6.8.0
Trino (PrestoSQL)	398	398	388	388
Zeppelin	0,1,1	0,1,1	0,1,1	0,1,1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notes de mise à jour 6.9.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.9.1. Les modifications ont été apportées à la version 6.9.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.9.1.

Modifications, améliorations et problèmes résolus

- Hadoop 3.3.3 a introduit une modification dans YARN (YARN-9608) qui maintient les nœuds sur lesquels les conteneurs s'exécutaient dans un état de mise hors service jusqu'à ce que l'application soit terminée. Cette modification permet de s'assurer que les données locales telles que les données réorganisées ne sont pas perdues et que vous n'avez pas besoin de réexécuter la tâche. Cette approche peut également entraîner une sous-utilisation des ressources sur les clusters avec ou sans activation de la mise à l'échelle gérée.

  Dans les versions 6.11.0 et supérieures d'Amazon EMR ainsi que dans les versions 6.8.1, 6.9.1 et 6.10.1, la valeur de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` est définie sur `false` dans `yarn-site.xml` pour résoudre ce problème.

  Bien que ce correctif règle les problèmes introduits par YARN-9608, il peut entraîner l'échec des tâches Hive en raison de la perte de données de réorganisation sur les clusters pour lesquels la mise à l'échelle gérée est activée. Nous avons atténué ce risque dans cette version en définissant également `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` pour les charges de travail Hive. Cette configuration n'est disponible qu'à partir de la version 6.11.0 d'Amazon EMR.
- Le collecteur de métriques n'enverra aucune métrique au plan de contrôle après le basculement du nœud primaire dans les clusters avec la configuration des groupes d'instances.
- Cette version élimine les nouvelles tentatives en cas d'échec des requêtes HTTP vers les points de terminaison du collecteur de métriques.

- Cette version inclut une modification qui permet aux clusters à haute disponibilité de se remettre d'un état défaillant après le redémarrage.
- Cette version corrige un problème selon lequel de grandes quantités créées par l'utilisateur UIDs provoquaient des exceptions de dépassement.
- Cette version corrige les problèmes de dépassement de délai avec le processus de reconfiguration d'Amazon EMR.
- Cette version inclut des correctifs de sécurité.
- Cette version résout un problème selon lequel les clusters exécutant des charges de travail sur Spark avec Amazon EMR peuvent recevoir silencieusement des résultats incorrects avec `contains`, `startsWith`, `endsWith` et `like`. Ce problème se produit lorsque vous utilisez les expressions sur des champs partitionnés contenant des métadonnées dans Amazon EMR Hive3 Metastore Server (HMS).
- Avec Amazon EMR 6.6.0 à 6.9.x, les requêtes INSERT avec partition dynamique et clause ORDER BY ou SORT BY auront toujours deux réducteurs. Ce problème est dû à la modification d'OSS HIVE-20703, qui place l'optimisation des partitions dynamiques de tri dans le cadre d'une décision basée sur les coûts. Si votre charge de travail ne nécessite pas le tri des partitions dynamiques, nous vous recommandons de définir la propriété `hive.optimize.sort.dynamic.partition.threshold` sur `-1` pour désactiver la nouvelle fonctionnalité et obtenir le nombre de réducteurs correctement calculé. Ce problème est résolu dans OSS Hive dans le cadre de HIVE-22269 et dans Amazon EMR 6.10.0.
- Hive peut subir des pertes de données lorsque vous utilisez HDFS comme répertoire scratch et que vous avez activé la fusion de petits fichiers, et que la table contient des chemins de partition statiques.
- Cette version corrige un problème de performance avec Hive si la fusion des petits fichiers (désactivée par défaut) est activée à la fin de la tâche ETL.
- Cette version corrige un problème de limitation du côté de Glue en l'absence de fonctions définies par l'utilisateur (UDF).
- Cette version corrige un problème qui supprime les journaux des conteneurs par le service d'agrégation des journaux des nœuds avant que le transmetteur de journaux ne puisse les envoyer vers S3 en cas de mise hors service de YARN.
- Cette version corrige la gestion des fichiers compactés/archivés avec un suivi permanent des fichiers de stockage pour. HBase

- Cette version corrige un problème qui affectait les performances de Spark lorsque vous définissez une valeur `true` par défaut pour la configuration `spark.yarn.heterogeneousExecutors.enabled` dans `spark-defaults.conf`.
- Cette version corrige un problème lié à l'échec de la lecture des données de shuffle par Reduce Task. Ce problème provoquait des échecs de requêtes Hive avec une erreur de mémoire corrompue.
- Cette version corrige un problème qui provoquait l'échec du fournisseur de nœuds si le service HDFS NameNode (NN) était bloqué en mode sécurisé lors du remplacement du nœud.
- Cette version ajoute un nouveau mécanisme de nouvelle tentative au flux de travail de dimensionnement des clusters pour les clusters EMR qui exécutent Presto ou Trino. Cette amélioration réduit le risque que le redimensionnement du cluster soit bloqué indéfiniment en raison de l'échec d'une seule opération de redimensionnement. Cela améliore également l'utilisation du cluster, car celui-ci augmente et diminue la capacité plus rapidement.
- Cette version améliore la logique de réduction de la taille des clusters afin que votre cluster ne tente pas une réduction d'échelle des nœuds principaux en dessous du paramètre de facteur de réplication HDFS défini pour le cluster. Cela répond à vos exigences en matière de redondance des données et réduit le risque de blocage d'une opération de dimensionnement.
- Le démon de gestion des journaux a été mis à niveau pour identifier tous les journaux en cours d'utilisation avec des descripteurs de fichiers ouverts sur le stockage d'instance local, ainsi que les processus associés. Cette mise à niveau garantit qu'Amazon EMR supprime correctement les fichiers et récupère de l'espace de stockage une fois les journaux archivés dans Amazon S3.
- Cette version inclut une amélioration du démon de gestion des journaux qui supprime les répertoires d'étapes vides et inutilisés dans le système de fichiers du cluster local. Un trop grand nombre de répertoires vides peut dégrader les performances des démons Amazon EMR et entraîner une surutilisation du disque.
- Cette version résout un problème qui peut survenir lorsque vous créez un nœud périphérique en répliquant l'un des nœuds primaires à partir d'un cluster comportant plusieurs nœuds primaires. Le nœud périphérique répliqué peut retarder les opérations de réduction d'échelle ou entraîner une utilisation élevée de la mémoire sur les nœuds primaires. Pour plus d'informations sur la création d'un nœud périphérique pour communiquer avec votre cluster EMR, consultez la section Edge Node Creator dans le `aws-samples` référentiel sur. GitHub
- Cette version améliore le processus d'automatisation utilisé par Amazon EMR pour remonter les volumes Amazon EBS sur une instance après un redémarrage.
- Cette version corrige un problème qui entraînait des écarts intermittents dans les métriques Hadoop publiées par Amazon EMR sur Amazon. CloudWatch

- Cette version résout un problème lié aux clusters EMR où une mise à jour du fichier de configuration YARN contenant la liste d'exclusion des nœuds du cluster est interrompue en raison d'une surutilisation du disque. La mise à jour incomplète entrave les futures opérations de réduction de la taille du cluster. Cette version garantit que votre cluster reste sain et que les opérations de dimensionnement fonctionnent comme prévu.
- Cette version améliore le démon de gestion des journaux sur le cluster afin de surveiller des dossiers de journaux supplémentaires dans votre cluster EMR. Cette amélioration permet de minimiser les scénarios de surutilisation des disques.
- Cette version redémarre automatiquement le démon de gestion des journaux sur le cluster lorsqu'il s'arrête. Cette amélioration réduit le risque que les nœuds apparaissent défectueux en raison d'une surutilisation du disque.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 822.0	4,1,4322	30 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

Versions des composants 6.9.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon

EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.2	Kit de développement logiciel Amazon SageMaker Spark
delta	2.1.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.3.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.6.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.23.0	Application de copie distribuée optimisée pour Amazon S3.

Composant	Version	Description
emr-s3-select	2.2.0	Connecteur S3Select EMR
emrfs	2,54,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.15.2	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.15.2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.3.3-amzn-1.1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.3.3-amzn-1.1	HDFS node-level service for storing blocks.

Composant	Version	Description
hadoop-hdfs-library	3.3.3-amzn-1.1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.3.3-amzn-1.1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.3.3-amzn-1.1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.3-amzn-1.1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.3-amzn-1.1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-1.1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.3-amzn-1.1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.3.3-amzn-1.1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.3.3-amzn-1.1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	2.4.13-amzn-0.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.13-amzn-0.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.13-amzn-0.1	HBase client en ligne de commande.
hbase-rest-server	2.4.13-amzn-0.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.13-amzn-0.1	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.13-amzn-0.1	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-2.1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-2.1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-2.1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	3.1.3-amzn-2.1	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-2.1	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-2.1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-2.1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.12.1-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.12.1-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.12.1-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.12.1-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.10.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
iceberg	0.14.1-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,7.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.2	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-connectors	6.0.0-SNAPSHOT	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.276-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.276-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.276-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	398-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	398-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	398-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.

Composant	Version	Description
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.3.0-amzn-1.1	Clients de ligne de commande Spark.
spark-history-server	3.3.0-amzn-1.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.3.0-amzn-1.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.3.0-amzn-1.1	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	22.08.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.10.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.

Composant	Version	Description
tez-on-yarn	0.10.2-amzn-0.1	Les bibliothèques et l'application tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 6.9.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.9.1

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.

Classifications	Description	Actions de reconfiguration
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Classifications	Description	Actions de reconfiguration
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.

Classifications	Description	Actions de reconfiguration
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.

Classifications	Description	Actions de reconfiguration
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.

Classifications	Description	Actions de reconfiguration
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

Journal des modifications 6.9.1

Journal des modifications pour la version 6.9.1 et les notes de mise à jour

Date	Événement	Description
30/08/2023	Mise à jour des notes de mise à jour	Ajout de plusieurs corrections liées au plan de contrôle dans les notes de mise à jour
21/08/2023	Publication de documents	Notes de mise à jour d'Amazon EMR 6.9.1 publiées pour la première fois
16/08/2023	Déploiement terminé	Amazon EMR 6.9.1 entièrement déployé dans toutes les régions prises en charge
04/08/2023	Première version	Amazon EMR 6.9.1 a été déployé pour la première fois dans des régions commerciales limitées

Amazon EMR version 6.9.0

Versions d’application 6.9.0

Cette version inclut les applications suivantes : Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.9.0	emr-6.8.1	emr-6.8.0	emr-6.7.0
AWS SDK pour Java	1,12,170	1,12,170	1,12,170	1,12,170
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,15	2,1,15	2,1,15	2,1,15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.1.0	-	-	-
Flink	1.15.2	1.15.1	1.15.1	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.13	2.4.12	2.4.12	2.4.4

	emr-6.9.0	emr-6.8.1	emr-6.8.0	emr-6.7.0
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.1-amzn-0	0.11.1-amzn-0	0.11.1-amzn-0	0.11.0-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.1-amzn-0	0.14.0-amzn-0	0.14.0-amzn-0	0.13.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,276	0,273	0,273	0,272
Spark	3.3.0	3.3.0	3.3.0	3.2.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.10.0	2.9.1	2.9.1	2.4.1
Tez	0,1,2	0.9.2	0.9.2	0.9.2

	emr-6.9.0	emr-6.8.1	emr-6.8.0	emr-6.7.0
Trino (PrestoSQL)	398	388	388	378
Zeppelin	0,1,1	0,1,1	0,1,1	0.10.0
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.7

Notes de mise à jour 6.9.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.9.0. Il s'agit des modifications apportées à la version Amazon EMR 6.8.0. Pour plus d'informations sur le calendrier de publication, consultez le journal des modifications.

Nouvelles fonctions

- La version 6.9.0 d'Amazon EMR prend en charge Apache Spark RAPIDS 22.08.0, Apache Hudi 0.12.1, Apache Iceberg 0.14.1, Trino 398 et Tez 0.10.2.
- La version 6.9.0 d'Amazon EMR inclut une nouvelle application open source, Delta Lake 2.1.0.
- L'intégration d'Amazon Redshift à Apache Spark est incluse dans les versions 6.9.0 et ultérieures d'Amazon EMR. Auparavant un outil open-source, l'intégration native est un connecteur Spark que vous pouvez utiliser pour créer des applications Apache Spark capables de lire et d'écrire des données sur Amazon Redshift et Amazon Redshift sans serveur. Pour de plus amples informations, veuillez consulter Utilisation de l'intégration d'Amazon Redshift pour Apache Spark avec Amazon EMR .
- La version 6.9.0 d'Amazon EMR ajoute la prise en charge de l'archivage des journaux dans Amazon S3 lors de la réduction de la taille du cluster. Auparavant, vous pouviez uniquement archiver les fichiers journaux sur Amazon S3 lors de la résiliation du cluster. Cette nouvelle fonctionnalité garantit que les fichiers journaux générés sur le cluster sont conservés sur Amazon S3 même après la résiliation du nœud. Pour plus d'informations, consultez Configuration de la journalisation et du débogage de cluster.
- Pour prendre en charge les requêtes de longue durée, Trino inclut désormais un mécanisme d'exécution tolérant aux pannes. L'exécution tolérante aux pannes atténue les échecs des requêtes en réessayant les requêtes qui ont échoué ou les tâches correspondantes.

- Vous pouvez utiliser Apache Flink sur Amazon EMR pour le traitement `BATCH` et `STREAM` unifié des tables Apache Hive ou des métadonnées de n'importe quelle source de tables Flink telle que Iceberg, Kinesis ou Kafka. Vous pouvez spécifier le catalogue de données AWS Glue comme métastore pour Flink à l'aide de l'API AWS Management Console, AWS CLI ou Amazon EMR. Pour de plus amples informations, veuillez consulter Configuration de Flink dans Amazon EMR.
- Vous pouvez désormais spécifier des rôles d'exécution AWS Identity and Access Management (IAM) et un contrôle d'accès AWS Lake Formation basé pour les requêtes Apache Spark, Apache Hive et Presto sur Amazon EMR sur des clusters EC2 avec Amazon AI Studio. SageMaker Pour obtenir des informations supplémentaires, consultez Configuration des rôles d'exécution pour les étapes d'Amazon EMR.

Problèmes connus

- Pour Amazon EMR version 6.9.0, Trino ne fonctionne pas sur les clusters activés pour Apache Ranger. Si vous devez utiliser Trino avec Ranger, contactez Support.
- Si vous utilisez l'intégration Amazon Redshift à Apache Spark et que vous disposez d'un champ de type heure, timetz, horodatage ou timestamptz avec une précision de l'ordre de la microseconde au format Parquet, le connecteur arrondit les valeurs temporelles à la milliseconde la plus proche. Pour contourner le problème, utilisez le paramètre `unload_s3_format` de format de déchargement du texte.
- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

- Les connexions aux clusters Amazon EMR depuis Amazon SageMaker AI Studio peuvent échouer par intermittence avec un code de réponse 403 Forbidden. Cette erreur se produit lorsque la configuration du rôle IAM sur le cluster prend plus de 60 secondes. Pour contourner le problème, vous pouvez installer un correctif Amazon EMR pour activer les nouvelles tentatives et augmenter le délai d'expiration à un minimum de 300 secondes. Suivez les étapes ci-dessous pour appliquer l'action d'amorçage lorsque vous lancez votre cluster.

  1. Téléchargez le script bootstrap et les fichiers RPM depuis l'Amazon S3 URIs suivant.

     ```
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/replace-rpms.sh
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/emr-secret-agent-1.18.0-SNAPSHOT20221121212949.noarch.rpm
     ```

  2. Chargez les fichiers de l'étape précédente sur un compartiment Amazon S3 dont vous êtes propriétaire. Le compartiment doit se trouver à l' Région AWS endroit où vous prévoyez de lancer le cluster.

  3. Incluez l'action d'amorçage suivante lorsque vous lancez votre cluster EMR. Remplacez *`bootstrap_URI`* et *`RPM_URI`* par le correspondant URIs provenant d'Amazon S3.

     ```
     --bootstrap-actions "Path=bootstrap_URI,Args=[RPM_URI]"
     ```

- Avec les versions 5.36.0 et 6.6.0 à 6.9.0 d'Amazon EMR, les composants de service `SecretAgent` et `RecordServer` peuvent subir une perte de données de journal en raison d'une configuration incorrecte du modèle de nom de fichier dans les propriétés de Log4j2. En cas de configuration incorrecte, les composants ne génèrent qu'un seul fichier journal par jour. Lorsque la stratégie de rotation est appliquée, elle remplace le fichier existant au lieu de générer un nouveau fichier journal comme prévu. Pour contourner le problème, utilisez une action d'amorçage pour générer des journaux toutes les heures et ajoutez un nombre entier auto-incrémenté dans le nom du fichier pour gérer la rotation.

  Pour les versions 6.6.0 à 6.9.0 d'Amazon EMR, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Pour Amazon EMR 5.36.0, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- Apache Flink fournit des FileSystem connecteurs natifs S3 FileSystem et Hadoop, qui permettent aux applications de créer FileSink et d'écrire les données dans Amazon S3. Cela FileSink échoue avec l'une des deux exceptions suivantes.

```
java.lang.UnsupportedOperationException: Recoverable writers on Hadoop are only
 supported for HDFS
```

```
Caused by: java.lang.NoSuchMethodError:
 org.apache.hadoop.io.retry.RetryPolicies.retryOtherThanRemoteAndSaslException(Lorg/
apache/hadoop/io/retry/RetryPolicy;Ljava/util/Map;)Lorg/apache/hadoop/io/retry/
RetryPolicy;
                                        at
 org.apache.hadoop.yarn.client.RMProxy.createRetryPolicy(RMProxy.java:302) ~[hadoop-
yarn-common-3.3.3-amzn-0.jar:?]
```

Pour contourner le problème, vous pouvez installer un correctif Amazon EMR, qui corrige le problème ci-dessus dans Flink. Suivez les étapes suivantes pour appliquer l'action d'amorçage lors du lancement de votre cluster.

1. Téléchargez le fichier flink-rpm dans votre compartiment Amazon S3. Votre chemin RPM est `s3://DOC-EXAMPLE-BUCKET/rpms/flink/`.
2. Téléchargez le script d'amorçage et les fichiers RPM depuis Amazon S3 en utilisant l'URI suivant. Remplacez *regionName* par l' Région AWS endroit où vous prévoyez de lancer le cluster.

   ```
   s3://emr-data-access-control-regionName/customer-bootstrap-actions/gcsc/replace-rpms.sh
   ```

3. Hadoop 3.3.3 a introduit une modification dans YARN (YARN-9608) qui maintient les nœuds sur lesquels les conteneurs s'exécutaient dans un état de mise hors service jusqu'à ce que

l'application soit terminée. Cette modification permet de s'assurer que les données locales telles que les données réorganisées ne sont pas perdues et que vous n'avez pas besoin de réexécuter la tâche. Dans Amazon EMR 6.8.0 et 6.9.0, cette approche peut également entraîner une sous-utilisation des ressources sur les clusters avec ou sans activation de la mise à l'échelle gérée.

Avec Amazon EMR 6.10.0, il existe une solution à ce problème qui consiste à définir la valeur de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` sur `false` dans `yarn-site.xml`. Dans les versions 6.11.0 et supérieures d'Amazon EMR, ainsi que 6.8.1, 6.9.1 et 6.10.1, la configuration est définie sur `false` par défaut pour résoudre ce problème.

Modifications, améliorations et problèmes résolus

- Pour Amazon EMR version 6.9.0 et versions ultérieures, tous les composants installés par Amazon EMR qui utilisent les bibliothèques Log4j utilisent Log4j version 2.17.1 ou ultérieure.
- Lorsque vous utilisez le connecteur DynamoDB avec Spark sur les versions 6.6.0, 6.7.0 et 6.8.0 d'Amazon EMR, toutes les lectures de votre table renvoient un résultat vide, même si la division d'entrée fait référence à des données non vides. La version 6.9.0 d'Amazon EMR résout ce problème.
- Amazon EMR 6.9.0 ajoute une prise en charge limitée du contrôle d'accès basé sur Lake Formation avec Apache Hudi lors de la lecture de données à l'aide de Spark SQL. La prise en charge concerne les requêtes SELECT utilisant Spark SQL et se limite au contrôle d'accès au niveau des colonnes. Pour plus d'informations, consultez Hudi et Lake Formation.
- Lorsque vous utilisez Amazon EMR 6.9.0 pour créer un cluster Hadoop avec les étiquettes de nœuds activées, l'API de métriques YARN renvoie des informations agrégées sur toutes les partitions, au lieu de la partition par défaut. Pour plus d'informations, consultez YARN-11414.
- Avec la version 6.9.0 d'Amazon EMR, nous avons mis à jour Trino vers la version 398, qui utilise Java 17. La version précédemment prise en charge de Trino pour Amazon EMR 6.8.0 était Trino 388 fonctionnant sous Java 11. Pour plus d'informations sur cette modification, consultez Mises à jour de Trino vers Java 17 sur le blog de Trino.
- Cette version corrige un problème d'incompatibilité de séquence temporelle entre Apache BigTop et Amazon EMR lors de la séquence de démarrage EC2 du cluster. Ce décalage se produit lorsqu'un système tente d'effectuer deux ou plusieurs opérations en même temps au lieu de les effectuer dans le bon ordre. Par conséquent, certaines configurations de cluster ont connu des délais de démarrage des instances et des temps de démarrage des clusters plus lents.

- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

> **Note**
>
> Cette version ne bénéficie plus de mises à jour automatiques de l'AMI puisqu'elle a été suivie d'une version supplémentaire de correctifs. La version du correctif est indiquée par le numéro qui suit la deuxième décimale (`6.8.1`). Pour savoir si vous utilisez la dernière version du correctif, consultez les versions disponibles dans le Guide des versions, ou consultez le menu déroulant des versions d'Amazon EMR lorsque vous créez un cluster dans la console, ou utilisez l'action d'API `ListReleaseLabels` ou de CLI `list-release-labels`. Pour être tenu au courant des nouvelles versions, abonnez-vous au flux RSS sur la page Quoi de neuf ?

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 418,0	4,14.311	3 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 404.1	4,14.311	18 avril 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0.2023 404.0	4,14.311	10 avril 2023	USA Est (Virginie du Nord), Europe (Paris)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 320,0	4,14,309	30 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 307.0	4,14,305	15 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 207.0	4,14,304	22 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2022 210.1	4,14,301	12 janvier 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 103.3	4,14.296	5 décembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

Versions des composants 6.9.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous

nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.2	Kit de développement logiciel Amazon SageMaker Spark
delta	2.1.0	Delta Lake est un format de table ouvert pour les jeux de données analytiques de grande taille.
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	3.3.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.6.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.23.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.2.0	Connecteur S3Select EMR

Composant	Version	Description
emrfs	2,54,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.15.2	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.15.2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.3.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.3.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.3.3-amzn-1	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	3.3.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.3.3-amzn-1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.3.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.3.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.3.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.3.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.3.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	2.4.13-amzn-0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.13-amzn-0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.13-amzn-0	HBase client en ligne de commande.
hbase-rest-server	2.4.13-amzn-0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.13-amzn-0	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.13-amzn-0	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	3.1.3-amzn-2	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-2	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.12.1-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.12.1-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.12.1-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.12.1-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.10.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
iceberg	0.14.1-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,7.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.2	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-connectors	6.0.0-SNAPSHOT	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	6.0.0	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.276-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.276-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.276-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	398-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	398-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	398-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.

Composant	Version	Description
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.3.0-amzn-1	Clients de ligne de commande Spark.
spark-history-server	3.3.0-amzn-1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.3.0-amzn-1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.3.0-amzn-1	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	22.08.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.10.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.

Composant	Version	Description
tez-on-yarn	0.10.2-amzn-0	Les bibliothèques et l'application tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 6.9.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.9.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.

Classifications	Description	Actions de reconfiguration
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
delta-defaults	Modification des valeurs dans le fichier delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j2	Modifiez les paramètres Livy log4j2.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Classifications	Description	Actions de reconfiguration
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Modifiez les valeurs dans le fichier lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-delta	Modification des valeurs dans le fichier delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Modification des valeurs dans le fichier exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.

Classifications	Description	Actions de reconfiguration
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.

Classifications	Description	Actions de reconfiguration
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.

Classifications	Description	Actions de reconfiguration
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

Journal des modifications 6.9.0

Journal des modifications pour la version 6.9.0 et les notes de mise à jour

Date	Événement	Description
30/08/2023	Mise à jour des notes de mise à jour	Ajout d'un correctif pour un problème de non-concordance des séquences de synchroni sation
21/08/2023	Mise à jour des notes de mise à jour	Ajout d'un problème connu avec Hadoop 3.3.3.
26/07/2023	Mettre à jour	Étiquettes `2.0.20230612.0` et `2.0.20230628.0` de la nouvelle version du système d'exploitation.
13/12	Mise à jour des notes de mise à jour	Fonctionnalité ajoutée et problème connu d'exécution avec l' SageMaker IA
29/11	Mise à jour des notes de mise à jour et de la documentation	Ajout d'une fonctionnalité pour l'intégration d'Amazon Redshift pour Apache Spark
23/11	Mise à jour des notes de mise à jour	Suppression de l'entrée Log4j

Date	Événement	Description
11/18	Déploiement terminé	Amazon EMR 6.9 entièrement déployé dans toutes les régions prises en charge
11/18	Publication de documents	Première publication des notes de mise à jour d'Amazon EMR 6.9
14/11	Première version	Amazon EMR 6.9 déployé dans des régions commerciales limitées

Amazon EMR version 6.8.1

Versions d’application 6.8.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
AWS SDK pour Java	1,12,170	1,12,170	1,12,170	1,12,170
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,15	2,1,15	2,1,15	2,1,10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.15.1	1.15.1	1.14.2	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.12	2.4.12	2.4.4	2.4.4
HCatalog	3.1.3	3.1.3	3.1.3	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.2
Hudi	0.11.1-amzn-0	0.11.1-amzn-0	0.11.0-amzn-0	0.10.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.0-amzn-0	0.14.0-amzn-0	0.13.1-amzn-0	0.13.1
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.8.0	1.8.0

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,273	0,273	0,272	0,267
Spark	3.3.0	3.3.0	3.2.1	3.2.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.9.1	2.9.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	388	388	378	367
Zeppelin	0,1,1	0,1,1	0.10.0	0.10.0
ZooKeeper	3.5.10	3.5.10	3.5.7	3.5.7

Notes de mise à jour 6.8.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.8.1. Les modifications ont été apportées à la version 6.8.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.8.1.

Modifications, améliorations et problèmes résolus

- Hadoop 3.3.3 a introduit une modification dans YARN (YARN-9608) qui maintient les nœuds sur lesquels les conteneurs s'exécutaient dans un état de mise hors service jusqu'à ce que l'application soit terminée. Cette modification permet de s'assurer que les données locales telles que les données réorganisées ne sont pas perdues et que vous n'avez pas besoin de réexécuter la tâche.

Cette approche peut également entraîner une sous-utilisation des ressources sur les clusters avec ou sans activation de la mise à l'échelle gérée.

Dans les versions 6.11.0 et supérieures d'Amazon EMR ainsi que dans les versions 6.8.1, 6.9.1 et 6.10.1, la valeur de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` est définie sur `false` dans `yarn-site.xml` pour résoudre ce problème.

Bien que ce correctif règle les problèmes introduits par YARN-9608, il peut entraîner l'échec des tâches Hive en raison de la perte de données de réorganisation sur les clusters pour lesquels la mise à l'échelle gérée est activée. Nous avons atténué ce risque dans cette version en définissant également `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` pour les charges de travail Hive. Cette configuration n'est disponible qu'à partir de la version 6.11.0 d'Amazon EMR.

- Le collecteur de métriques n'enverra aucune métrique au plan de contrôle après le basculement du nœud primaire dans les clusters avec la configuration des groupes d'instances.
- Cette version élimine les nouvelles tentatives en cas d'échec des requêtes HTTP vers les points de terminaison du collecteur de métriques.
- Cette version inclut une modification qui permet aux clusters à haute disponibilité de se remettre d'un état défaillant après le redémarrage.
- Cette version corrige un problème selon lequel de grandes quantités créées par l'utilisateur UIDs provoquaient des exceptions de dépassement.
- Cette version corrige les problèmes de dépassement de délai avec le processus de reconfiguration d'Amazon EMR.
- Cette version permet d'éviter un problème selon lequel l'échec de la reconfiguration pourrait interrompre d'autres processus indépendants.
- Cette version inclut des correctifs de sécurité.
- Cette version résout un problème selon lequel les clusters exécutant des charges de travail sur Spark avec Amazon EMR peuvent recevoir silencieusement des résultats incorrects avec `contains`, `startsWith`, `endsWith` et `like`. Ce problème se produit lorsque vous utilisez les expressions sur des champs partitionnés contenant des métadonnées dans Amazon EMR Hive3 Metastore Server (HMS).
- Avec Amazon EMR 6.6.0 à 6.9.x, les requêtes INSERT avec partition dynamique et clause ORDER BY ou SORT BY auront toujours deux réducteurs. Ce problème est dû à la modification d'OSS HIVE-20703, qui place l'optimisation des partitions dynamiques de tri dans le cadre d'une décision basée sur les coûts. Si votre charge de travail ne nécessite

pas le tri des partitions dynamiques, nous vous recommandons de définir la propriété `hive.optimize.sort.dynamic.partition.threshold` sur `-1` pour désactiver la nouvelle fonctionnalité et obtenir le nombre de réducteurs correctement calculé. Ce problème est résolu dans OSS Hive dans le cadre de HIVE-22269 et dans Amazon EMR 6.10.0.

- Hive peut subir des pertes de données lorsque vous utilisez HDFS comme répertoire scratch et que vous avez activé la fusion de petits fichiers, et que la table contient des chemins de partition statiques.
- Cette version corrige un problème de performance avec Hive si la fusion des petits fichiers (désactivée par défaut) est activée à la fin de la tâche ETL.
- Cette version corrige un problème de limitation du côté de Glue en l'absence de fonctions définies par l'utilisateur (UDF).
- Cette version corrige un problème qui supprime les journaux des conteneurs par le service d'agrégation des journaux des nœuds avant que le transmetteur de journaux ne puisse les envoyer vers S3 en cas de mise hors service de YARN.
- Cette version corrige la gestion des fichiers compactés/archivés avec un suivi permanent des fichiers de stockage pour. HBase
- Cette version corrige un problème qui affectait les performances de Spark lorsque vous définissez une valeur `true` par défaut pour la configuration `spark.yarn.heterogeneousExecutors.enabled` dans `spark-defaults.conf`.
- Cette version corrige un problème lié à l'échec de la lecture des données de shuffle par Reduce Task. Ce problème provoquait des échecs de requêtes Hive avec une erreur de mémoire corrompue.
- Cette version corrige un problème qui provoquait l'échec du fournisseur de nœuds si le service HDFS NameNode (NN) était bloqué en mode sécurisé lors du remplacement du nœud.
- Cette version ajoute un nouveau mécanisme de nouvelle tentative au flux de travail de dimensionnement des clusters pour les clusters EMR qui exécutent Presto ou Trino. Cette amélioration réduit le risque que le redimensionnement du cluster soit bloqué indéfiniment en raison de l'échec d'une seule opération de redimensionnement. Cela améliore également l'utilisation du cluster, car celui-ci augmente et diminue la capacité plus rapidement.
- Cette version améliore la logique de réduction de la taille des clusters afin que votre cluster ne tente pas une réduction d'échelle des nœuds principaux en dessous du paramètre de facteur de réplication HDFS défini pour le cluster. Cela répond à vos exigences en matière de redondance des données et réduit le risque de blocage d'une opération de dimensionnement.

- Le démon de gestion des journaux a été mis à niveau pour identifier tous les journaux en cours d'utilisation avec des descripteurs de fichiers ouverts sur le stockage d'instance local, ainsi que les processus associés. Cette mise à niveau garantit qu'Amazon EMR supprime correctement les fichiers et récupère de l'espace de stockage une fois les journaux archivés dans Amazon S3.
- Cette version inclut une amélioration du démon de gestion des journaux qui supprime les répertoires d'étapes vides et inutilisés dans le système de fichiers du cluster local. Un trop grand nombre de répertoires vides peut dégrader les performances des démons Amazon EMR et entraîner une surutilisation du disque.
- Cette version résout un problème qui peut survenir lorsque vous créez un nœud périphérique en répliquant l'un des nœuds primaires à partir d'un cluster comportant plusieurs nœuds primaires. Le nœud périphérique répliqué peut retarder les opérations de réduction d'échelle ou entraîner une utilisation élevée de la mémoire sur les nœuds primaires. Pour plus d'informations sur la création d'un nœud périphérique pour communiquer avec votre cluster EMR, consultez la section Edge Node Creator dans le `aws-samples` référentiel sur. GitHub
- Cette version améliore le processus d'automatisation utilisé par Amazon EMR pour remonter les volumes Amazon EBS sur une instance après un redémarrage.
- Cette version corrige un problème qui entraînait des écarts intermittents dans les métriques Hadoop publiées par Amazon EMR sur Amazon. CloudWatch
- Cette version résout un problème lié aux clusters EMR où une mise à jour du fichier de configuration YARN contenant la liste d'exclusion des nœuds du cluster est interrompue en raison d'une surutilisation du disque. La mise à jour incomplète entrave les futures opérations de réduction de la taille du cluster. Cette version garantit que votre cluster reste sain et que les opérations de dimensionnement fonctionnent comme prévu.
- Cette version améliore le démon de gestion des journaux sur le cluster afin de surveiller des dossiers de journaux supplémentaires dans votre cluster EMR. Cette amélioration permet de minimiser les scénarios de surutilisation des disques.
- Cette version redémarre automatiquement le démon de gestion des journaux sur le cluster lorsqu'il s'arrête. Cette amélioration réduit le risque que les nœuds apparaissent défectueux en raison d'une surutilisation du disque.
- Cette version ajoute la prise en charge de l'archivage des journaux sur Amazon S3 lors de la réduction de la capacité du cluster. Auparavant, vous pouviez uniquement archiver les fichiers journaux sur Amazon S3 lors de la résiliation du cluster. Cette nouvelle fonctionnalité garantit que les fichiers journaux générés sur le cluster sont conservés sur Amazon S3 même après la résiliation du nœud. Pour plus d'informations, consultez Configuration de la journalisation et du débogage de cluster.

- Cette version corrige un problème qui s'est produit lorsque l'URI Amazon S3 pour une action d'amorçage se terminait par un numéro de port, par exemple : `a.b.c.d:4345`. Amazon EMR les analysait incorrectement URIs, de sorte que toutes les actions d'amorçage associées échoueraient.
- Cette version corrige un problème d'incompatibilité de séquence temporelle entre Apache BigTop et Amazon EMR lors de la séquence de démarrage EC2 du cluster. Ce décalage se produit lorsqu'un système tente d'effectuer deux ou plusieurs opérations en même temps au lieu de les effectuer dans le bon ordre. Par conséquent, certaines configurations de cluster ont connu des délais de démarrage des instances et des temps de démarrage des clusters plus lents.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), AWS GovCloud (USA Est), AWS GovCloud (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), Chine (Pékin), Chine (Ningxia), Moyen-Orient (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv),
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (centre), Israël (Tel Aviv), (États-Unis ouest) AWS GovCloud , (États-Un is est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e) Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 822.0	4,1,4322	30 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e) Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e) Afrique (Le Cap), Amérique du Sud (São Paulo), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

Versions des composants 6.8.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon

EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.2	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.2.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.22.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.1.0	Connecteur S3Select EMR
emrfs	2,53,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.

Composant	Version	Description
flink-client	1.15.1	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.15.1	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.2.1-amzn-8.1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-8.1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-8.1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.2.1-amzn-8.1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	3.2.1-amzn-8.1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.2.1-amzn-8.1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-8.1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-8.1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.2.1-amzn-8.1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.2.1-amzn-8.1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-8.1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.12-amzn-0.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	2.4.12-amzn-0.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.12-amzn-0.1	HBase client en ligne de commande.
hbase-rest-server	2.4.12-amzn-0.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.12-amzn-0.1	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.12-amzn-0.1	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-1.1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-1.1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-1.1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-1.1	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-1.1	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	3.1.3-amzn-1.1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-1.1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.11.1-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.11.1-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.11.1-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.11.1-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.10.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	0.14.0-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.

Composant	Version	Description
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,7.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.2	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	5.1.2	Connecteurs Apache Phoenix pour Spark-3

Composant	Version	Description
phoenix-query-server	5.1.2	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.273.3-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.273.3-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.273.3-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	388-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	388-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	388-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.

Composant	Version	Description
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.3.0-amzn-0.1	Clients de ligne de commande Spark.
spark-history-server	3.3.0-amzn-0.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.3.0-amzn-0.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.3.0-amzn-0.1	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	22.06.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.9.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 6.8.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.8.1

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.

Classifications	Description	Actions de reconfiguration
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.

Classifications	Description	Actions de reconfiguration
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Journal des modifications 6.8.1

Journal des modifications pour la version 6.8.1 et les notes de mise à jour

Date	Événement	Description
30/08/2023	Mise à jour des notes de mise à jour	Ajout de plusieurs corrections liées au plan de contrôle dans les notes de mise à jour
21/08/2023	Publication de documents	Notes de mise à jour d'Amazon EMR 6.8.1 publiées pour la première fois
16/08/2023	Déploiement terminé	Amazon EMR 6.8.1 entièrement déployé dans toutes les régions prises en charge
04/08/2023	Première version	Amazon EMR 6.8.1 a été déployé pour la première fois dans des régions commerciales limitées

Amazon EMR version 6.8.0

Versions d’application 6.8.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR

- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.8.0	emr-6.7.0	emr-6.6.0	emr-6.5.0
AWS SDK pour Java	1,12,170	1,12,170	1,12,170	1,12,31
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,15	2,1,15	2,1,10	2,1,10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.15.1	1.14.2	1.14.2	1.14.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.12	2.4.4	2.4.4	2.4.4
HCatalog	3.1.3	3.1.3	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.2	3.1.2
Hudi	0.11.1-amzn-0	0.11.0-amzn-0	0.10.1-amzn-0	0.9.0-amzn-1
Hue	4.10.0	4.10.0	4.10.0	4.9.0
Iceberg	0.14.0-amzn-0	0.13.1-amzn-0	0.13.1	0.12.0
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1

	emr-6.8.0	emr-6.7.0	emr-6.6.0	emr-6.5.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.8.0	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,273	0,272	0,267	0,261
Spark	3.3.0	3.2.1	3.2.0	3.1.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.9.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	388	378	367	360
Zeppelin	0,1,1	0.10.0	0.10.0	0.10.0
ZooKeeper	3.5.10	3.5.7	3.5.7	3.5.7

Notes de mise à jour 6.8.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.8.0. Les modifications ont été apportées à la version 6.7.0.

Nouvelles fonctions

- La fonctionnalité Amazon EMR Steps prend désormais en charge le point de terminaison et les clients Apache Livy. JDBC/ODBC Pour obtenir des informations supplémentaires, consultez Configuration des rôles d'exécution pour les étapes d'Amazon EMR.

- La version 6.8.0 d'Amazon EMR est fournie avec la version 2.4.12 d'Apache HBase . Avec cette HBase version, vous pouvez à la fois archiver et supprimer vos HBase tables. Le processus d'archivage Amazon S3 renomme tous les fichiers de table dans le répertoire d'archive. Ce processus peut être long et coûteux. Vous pouvez désormais ignorer le processus d'archivage et supprimer rapidement des tables volumineuses. Pour de plus amples informations, veuillez consulter Utilisation de la HBase coque.

Problèmes connus

- Hadoop 3.3.3 a introduit une modification dans YARN (YARN-9608) qui maintient les nœuds sur lesquels les conteneurs s'exécutaient dans un état de mise hors service jusqu'à ce que l'application soit terminée. Cette modification permet de s'assurer que les données locales telles que les données réorganisées ne sont pas perdues et que vous n'avez pas besoin de réexécuter la tâche. Dans Amazon EMR 6.8.0 et 6.9.0, cette approche peut également entraîner une sous-utilisation des ressources sur les clusters avec ou sans activation de la mise à l'échelle gérée.

  Avec Amazon EMR 6.10.0, il existe une solution à ce problème qui consiste à définir la valeur de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` sur `false` dans `yarn-site.xml`. Dans les versions 6.11.0 et supérieures d'Amazon EMR, ainsi que 6.8.1, 6.9.1 et 6.10.1, la configuration est définie sur `false` par défaut pour résoudre ce problème.

Modifications, améliorations et problèmes résolus

- Lorsqu'Amazon EMR version 6.5.0, 6.6.0 ou 6.7.0 lisait les tables Apache Phoenix via le shell Apache Spark, Amazon EMR a produit une erreur `NoSuchMethodError`. La version 6.8.0 d'Amazon EMR résout ce problème.
- La version 6.8.0 d'Amazon EMR est fournie avec Apache Hudi 0.11.1 ; toutefois, les clusters Amazon EMR 6.8.0 sont également compatibles avec le `hudi-spark3.3-bundle_2.12` open source de Hudi 0.12.0.
- La version 6.8.0 d'Amazon EMR est fournie avec la version 3.3.0 d'Apache Spark. Cette version de Spark utilise Apache Log4j 2 et le fichier `log4j2.properties` pour configurer Log4j dans les processus Spark. Si vous utilisez Spark dans le cluster ou si vous créez des clusters EMR avec des paramètres de configuration personnalisés, et que vous voulez passer à la version 6.8.0 d'Amazon EMR, vous devez migrer vers la nouvelle classification de configuration `spark-log4j2`

et le nouveau format de clé pour Apache Log4j 2. Pour de plus amples informations, veuillez consulter Migration d'Apache Log4j 1.x vers Log4j 2.x.

- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

> **Note**
>
> Cette version ne bénéficie plus de mises à jour automatiques de l'AMI puisqu'elle a été suivie d'une version supplémentaire de correctifs. La version du correctif est indiquée par le numéro qui suit la deuxième décimale (`6.8.1`). Pour savoir si vous utilisez la dernière version du correctif, consultez les versions disponibles dans le Guide des versions, ou consultez le menu déroulant des versions d'Amazon EMR lorsque vous créez un cluster dans la console, ou utilisez l'action d'API `ListReleaseLabels` ou de CLI `list-release-labels`. Pour être tenu au courant des nouvelles versions, abonnez-vous au flux RSS sur la page Quoi de neuf ?

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), AWS GovCloud (USA Est), AWS GovCloud (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv)
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 418,0	4,14.311	3 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 404.1	4,14.311	18 avril 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0.2023 404.0	4,14.311	10 avril 2023	USA Est (Virginie du Nord), Europe (Paris)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 320,0	4,14,309	30 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 307.0	4,14,305	15 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 207.0	4,14,304	22 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 119.1	4,14,301	3 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2022 210.1	4,14,301	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 103.3	4,14.296	5 décembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 004.0	4,14.294	2 novembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 912.1	4,14,291	6 septembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

Problèmes connus

- Lorsque vous utilisez le connecteur DynamoDB avec Spark sur les versions 6.6.0, 6.7.0 et 6.8.0 d'Amazon EMR, toutes les lectures de votre table renvoient un résultat vide, même si la division d'entrée fait référence à des données non vides. Cela est dû au fait que Spark 3.2.0 définit `spark.hadoopRDD.ignoreEmptySplits` sur `true` par défaut. Pour contourner le problème,

définissez explicitement `spark.hadoopRDD.ignoreEmptySplits` sur `false`. La version 6.9.0 d'Amazon EMR résout ce problème.

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.
- Avec les versions 5.36.0 et 6.6.0 à 6.9.0 d'Amazon EMR, les composants de service `SecretAgent` et `RecordServer` peuvent subir une perte de données de journal en raison d'une configuration incorrecte du modèle de nom de fichier dans les propriétés de Log4j2. En cas de configuration incorrecte, les composants ne génèrent qu'un seul fichier journal par jour. Lorsque la stratégie de rotation est appliquée, elle remplace le fichier existant au lieu de générer un nouveau fichier journal comme prévu. Pour contourner le problème, utilisez une action d'amorçage pour générer des journaux toutes les heures et ajoutez un nombre entier auto-incrémenté dans le nom du fichier pour gérer la rotation.

  Pour les versions 6.6.0 à 6.9.0 d'Amazon EMR, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
  ```

  Pour Amazon EMR 5.36.0, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Pour plus d'informations sur le calendrier de publication, consultez le journal des modifications.

Versions des composants 6.8.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.2	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	3.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.7.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.22.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.1.0	Connecteur S3Select EMR
emrfs	2,53,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.15.1	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.15.1	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	3.2.1-amzn-8	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-8	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-8	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.2.1-amzn-8	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-8	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.2.1-amzn-8	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-8	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-8	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.2.1-amzn-8	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	3.2.1-amzn-8	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-8	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.12-amzn-0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.12-amzn-0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.12-amzn-0	HBase client en ligne de commande.
hbase-rest-server	2.4.12-amzn-0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.12-amzn-0	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.12-amzn-0	Outil de réparation pour les HBase clusters Apache.
hcatalog-client	3.1.3-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	3.1.3-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-1	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-1	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.11.1-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.11.1-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.11.1-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.

Composant	Version	Description
hudi-spark	0.11.1-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.10.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
iceberg	0.14.0-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.9.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,7.0	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.2	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	5.1.2	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	5.1.2	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.273.3-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.273.3-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.273.3-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	388-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.

Composant	Version	Description
trino-worker	388-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	388-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.3.0-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.3.0-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.3.0-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.3.0-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	22.06.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.

Composant	Version	Description
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.9.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0,1,1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.5.10	ZooKeeper client en ligne de commande.

Classifications des configurations 6.8.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions

de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.8.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode,

Classifications	Description	Actions de reconfiguration
		SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.

Classifications	Description	Actions de reconfiguration
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j2	Modifiez les valeurs dans le fichier log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.

Classifications	Description	Actions de reconfiguration
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Journal des modifications 6.8.0

Journal des modifications pour la version 6.8.0 et les notes de mise à jour

Date	Événement	Description
21/08/2023	Mettre à jour	Ajout d'un problème connu avec Hadoop 3.3.3.
26/07/2023	Mettre à jour	Étiquettes `2.0.20230612.0` et `2.0.20230628.0` de la nouvelle version du système d'exploitation.
05.09-06	Déploiement terminé	Amazon EMR 6.8 entièrement déployé dans toutes les régions prises en charge
05.09-06	Publication initiale	Première publication des notes de mise à jour d'Amazon EMR 6.8
08-31	Première version	Amazon EMR 6.8 déployé dans des régions commerciales limitées

Amazon EMR version 6.7.0

Versions d’application 6.7.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.7.0	emr-6.6.0	emr-6.5.0	emr-6.4.0
AWS SDK pour Java	1,12,170	1,12,170	1,12,31	1,12,31
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,15	2,1,10	2,1,10	2,1,10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.0	1.13.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.4.4	2.4.4
HCatalog	3.1.3	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.2	3.1.2	3.1.2
Hudi	0.11.0-amzn-0	0.10.1-amzn-0	0.9.0-amzn-1	0.8.0-amzn-0
Hue	4.10.0	4.10.0	4.9.0	4.9.0
Iceberg	0.13.1-amzn-0	0.13.1	0.12.0	-

	emr-6.7.0	emr-6.6.0	emr-6.5.0	emr-6.4.0
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,272	0,267	0,261	0,254,1
Spark	3.2.1	3.2.0	3.1.2	3.1.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	378	367	360	359
Zeppelin	0.10.0	0.10.0	0.10.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.5.7	3.5.7

Notes de mise à jour 6.7.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.7.0. Les modifications ont été apportées à la version 6.6.0.

Date de parution initiale : 15 juillet 2022

Nouvelles fonctions

- Amazon EMR prend désormais en charge Apache Spark 3.2.1, Apache Hive 3.1.3, HUDI 0.11, PrestoDB 0.272 et Trino 0.378.
- Prend en charge les contrôles d'accès basés sur les rôles IAM et les formations lacustres avec des étapes EMR (Spark, Hive) pour Amazon EMR sur des clusters. EC2
- Prend en charge les instructions de définition de données Apache Spark sur les clusters compatibles avec Apache Ranger. Cela inclut désormais la prise en charge des applications Trino lisant et écrivant des métadonnées Apache Hive sur des clusters compatibles avec Apache Ranger. Pour plus d'informations, consultez Mise en place d'une gouvernance fédérée à l'aide de Trino et d'Apache Ranger sur Amazon EMR.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong),

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Asie-Pacifique (Hyderaba d), USA Est (N. Virginie) , Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (Singapour), Asie-Pacifique (Sydney) AWS GovCloud), Israël (Tel Aviv), Canada (centre), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Émirats arabes unis), Canada (Centre), (États-Unis- AWS GovCloud Est), AWS GovCloud (US-Ouest), Chine (Pékin),

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Israël (Tel Aviv), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Espagne), Europe (Zurich)
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 822.0	4,1,4322	30 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 418,0	4,14.311	3 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 404.1	4,14.311	18 avril 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0.2023 404.0	4,14.311	10 avril 2023	USA Est (Virginie du Nord), Europe (Paris)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 320,0	4,14,309	30 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 307.0	4,14,305	15 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 207.0	4,14,304	22 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 119.1	4,14,301	3 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2022 210.1	4,14,301	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 103.3	4,14.296	5 décembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 004.0	4,14.294	2 novembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 912.1	4,14,291	7 octobre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0,2022 719,0	4,14.287	10 août 2022	`us-west-1` , `eu-west-3` , `eu-north-1` , `ap-south-1` , `me-south-1`

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 606.1	4,14,281	15 juillet 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

Problèmes connus

- Lorsque les versions 6.5.0, 6.6.0 ou 6.7.0 d'Amazon EMR lisent les tables Apache Phoenix via le shell Apache Spark, une erreur `NoSuchMethodError` se produit car Amazon EMR utilise une `Hbase.compat.version` incorrecte. La version 6.8.0 d'Amazon EMR résout ce problème.
- Lorsque vous utilisez le connecteur DynamoDB avec Spark sur les versions 6.6.0, 6.7.0 et 6.8.0 d'Amazon EMR, toutes les lectures de votre table renvoient un résultat vide, même si la division

d'entrée fait référence à des données non vides. Cela est dû au fait que Spark 3.2.0 définit `spark.hadoopRDD.ignoreEmptySplits` sur `true` par défaut. Pour contourner le problème, définissez explicitement `spark.hadoopRDD.ignoreEmptySplits` sur `false`. La version 6.9.0 d'Amazon EMR résout ce problème.

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.
- Avec les versions 5.36.0 et 6.6.0 à 6.9.0 d'Amazon EMR, les composants de service `SecretAgent` et `RecordServer` peuvent subir une perte de données de journal en raison d'une configuration incorrecte du modèle de nom de fichier dans les propriétés de Log4j2. En cas de configuration incorrecte, les composants ne génèrent qu'un seul fichier journal par jour. Lorsque la stratégie de rotation est appliquée, elle remplace le fichier existant au lieu de générer un nouveau fichier journal comme prévu. Pour contourner le problème, utilisez une action d'amorçage pour générer des journaux toutes les heures et ajoutez un nombre entier auto-incrémenté dans le nom du fichier pour gérer la rotation.

  Pour les versions 6.6.0 à 6.9.0 d'Amazon EMR, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Pour Amazon EMR 5.36.0, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- L'API `GetClusterSessionCredentials` n'est pas prise en charge avec les clusters qui s'exécutent sur Amazon EMR 6.7 ou version antérieure.

Versions des composants 6.7.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	3.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.6.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.22.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.1.0	Connecteur S3Select EMR
emrfs	2,52,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.14.2	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.14.2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.2.1-amzn-7	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-7	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-7	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.2.1-amzn-7	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-7	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.2.1-amzn-7	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-7	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-7	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	3.2.1-amzn-7	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.2.1-amzn-7	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-7	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.4-amzn-3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.4-amzn-3	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.4-amzn-3	HBase client en ligne de commande.
hbase-rest-server	2.4.4-amzn-3	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.4-amzn-3	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.4-amzn-3	Outil de réparation pour les HBase clusters Apache.

Composant	Version	Description
hcatalog-client	3.1.3-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.3-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.3-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.3-amzn-0	Client de ligne de commande Hive.
hive-hbase	3.1.3-amzn-0	Client Hive-hbase.
hive-metastore-server	3.1.3-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.3-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.11.0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.11.0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.

Composant	Version	Description
hudi-trino	0.11.0-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.11.0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.10.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
iceberg	0.13.1-amzn-0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.8.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,0,194	Pilotes Nvidia et boîte à outils Cuda

Composant	Version	Description
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.2	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	5.1.2	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	5.1.2	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.272-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.272-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.272-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.

Composant	Version	Description
trino-coordinator	378-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	378-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	378-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.2.1-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.2.1-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.2.1-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.2.1-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.

Composant	Version	Description
spark-rapids	22.02.0-amzn-1	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0.10.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.7	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.5.7	ZooKeeper client en ligne de commande.

Classifications des configurations 6.7.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.7.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.

Classifications	Description	Actions de reconfiguration
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.

Classifications	Description	Actions de reconfiguration
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Amazon EMR version 6.6.0

Versions d’application 6.6.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.6.0	emr-6.5.0	emr-6.4.0	emr-6.3.1
AWS SDK pour Java	1,12,170	1,12,31	1,12,31	1,1,977
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,10	2,1,10	2,1,10	2,1,10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.0	1.13.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.4.4	2.2.6
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.10.1-amzn-0	0.9.0-amzn-1	0.8.0-amzn-0	0.7.0-amzn-0

	emr-6.6.0	emr-6.5.0	emr-6.4.0	emr-6.3.1
Hue	4.10.0	4.9.0	4.9.0	4.9.0
Iceberg	0.13.1	0.12.0	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.2.2
Livy	0.7.1	0.7.1	0.7.1	0.7.0
MXNet	1.8.0	1.8.0	1.8.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,267	0,261	0,254,1	0,245,1
Spark	3.2.0	3.1.2	3.1.2	3.1.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	367	360	359	350
Zeppelin	0.10.0	0.10.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.5.7	3.4.14

Notes de mise à jour 6.6.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.6.0. Les modifications ont été apportées à la version 6.5.0.

Date de parution initiale : 9 mai 2022

Date de mise à jour de la documentation : 15 juin 2022

Nouvelles fonctions

- Amazon EMR 6.6 prend désormais en charge Apache Spark 3.2, Apache Spark RAPIDS 22.02, CUDA 11, Apache Hudi 0.10.1, Apache Iceberg 0.13, Trino 0.367 et PrestoDB 0.267.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), USA Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Osaka), Europe (Londres) , Europe (Milan), Asie-Paci fique (Tokyo), (États-Un is est) AWS GovCloud , (États-Unis ouest), ouest des États-Unis (Oregon), ouest des États-Unis (Californie du Nord), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Canada (centre), Chine (Ningxia) AWS GovCloud

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 822.0	4,1,4322	30 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 418,0	4,14.311	3 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 404.1	4,14.311	18 avril 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0.2023 404.0	4,14.311	10 avril 2023	USA Est (Virginie du Nord), Europe (Paris)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 320,0	4,14,309	30 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 307.0	4,14,305	15 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 207.0	4,14,304	22 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 119.1	4,14,301	3 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2022 210.1	4,14,301	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 103.3	4,14.296	5 décembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 004.0	4,14.294	2 novembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 912.1	4,14,291	7 octobre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0,2022 805.0	4,14.287	30 août 2022	`us-west-1`

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2022 719,0	4,14.287	10 août 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 426.0	4,14,281	10 juin 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 406.1	4,14,275	2 mai 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

- Avec Amazon EMR 6.6 et les versions ultérieures, les applications qui utilisent Log4j 1.x et Log4j 2.x sont mises à niveau pour utiliser respectivement Log4j 1.2.17 (ou supérieur) et Log4j 2.17.1 (ou supérieur), et n'ont pas besoin d'utiliser les actions d'amorçage fournies pour atténuer les problèmes liés aux CVE.
- [Mise à l'échelle gérée] Optimisation de la mise à l'échelle gérée des données de réorganisation Spark – Pour Amazon EMR versions 5.34.0 et ultérieures, et EMR versions 6.4.0 et ultérieures, la mise à l'échelle gérée prend désormais en compte les données de réorganisation Spark (données que Spark redistribue entre les partitions pour effectuer des opérations spécifiques).

Pour plus d'informations sur les opérations de réorganisation, consultez Utilisation de la mise à l'échelle gérée par EMR dans Amazon EMR dans le Guide de gestion Amazon EMR et le Guide de programmation Spark.

- À partir d'Amazon EMR 5.32.0 et 6.5.0, le dimensionnement dynamique de l'exécuteur pour Apache Spark est activé par défaut. Pour activer ou désactiver cette fonctionnalité, vous pouvez utiliser le paramètre de configuration `spark.yarn.heterogeneousExecutors.enabled`.

Modifications, améliorations et problèmes résolus

- Amazon EMR réduit le temps de démarrage des clusters de 80 secondes en moyenne pour les clusters qui utilisent l'option AMI par défaut d'EMR et n'installent que des applications courantes, telles qu'Apache Hadoop, Apache Spark et Apache Hive.

Problèmes connus

- Lorsque les versions 6.5.0, 6.6.0 ou 6.7.0 d'Amazon EMR lisent les tables Apache Phoenix via le shell Apache Spark, une erreur `NoSuchMethodError` se produit car Amazon EMR utilise une `Hbase.compat.version` incorrecte. La version 6.8.0 d'Amazon EMR résout ce problème.
- Lorsque vous utilisez le connecteur DynamoDB avec Spark sur les versions 6.6.0, 6.7.0 et 6.8.0 d'Amazon EMR, toutes les lectures de votre table renvoient un résultat vide, même si la division d'entrée fait référence à des données non vides. Cela est dû au fait que Spark 3.2.0 définit `spark.hadoopRDD.ignoreEmptySplits` sur `true` par défaut. Pour contourner le problème, définissez explicitement `spark.hadoopRDD.ignoreEmptySplits` sur `false`. La version 6.9.0 d'Amazon EMR résout ce problème.
- Sur les clusters de longue durée de Trino, Amazon EMR 6.6.0 active les paramètres de journalisation du récupérateur de mémoire dans le fichier jvm.config de Trino afin d'obtenir de meilleures informations à partir des journaux du récupérateur de mémoire. Cette modification ajoute de nombreux journaux de collecte des déchets au fichier launcher.log (/var/log/trino/launcher.log). Si vous utilisez des clusters Trino dans Amazon EMR 6.6.0, vous pouvez rencontrer des nœuds à court d'espace disque après quelques jours d'exécution du cluster en raison des journaux ajoutés.

  La solution à ce problème consiste à exécuter le script ci-dessous en tant qu'action d'amorçage afin de désactiver les paramètres de journalisation du récupérateur de mémoire dans jvm.config lors de la création ou du clonage du cluster pour Amazon EMR 6.6.0.

```
#!/bin/bash
  set -ex
  PRESTO_PUPPET_DIR='/var/aws/emr/bigtop-deploy/puppet/modules/trino'
  sudo bash -c "sed -i '/-Xlog/d' ${PRESTO_PUPPET_DIR}/templates/jvm.config"
```

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.
- Avec les versions 5.36.0 et 6.6.0 à 6.9.0 d'Amazon EMR, les composants de service `SecretAgent` et `RecordServer` peuvent subir une perte de données de journal en raison d'une configuration incorrecte du modèle de nom de fichier dans les propriétés de Log4j2. En cas de configuration incorrecte, les composants ne génèrent qu'un seul fichier journal par jour. Lorsque la stratégie de rotation est appliquée, elle remplace le fichier existant au lieu de générer un nouveau fichier journal comme prévu. Pour contourner le problème, utilisez une action d'amorçage pour générer des journaux toutes les heures et ajoutez un nombre entier auto-incrémenté dans le nom du fichier pour gérer la rotation.

  Pour les versions 6.6.0 à 6.9.0 d'Amazon EMR, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Pour Amazon EMR 5.36.0, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Versions des composants 6.6.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	3.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.5.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.20.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.1.0	Connecteur S3Select EMR
emrfs	2,50,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.14.2	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.14.2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.2.1-amzn-6	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-6	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-6	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.2.1-amzn-6	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-6	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.2.1-amzn-6	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-6	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-6	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	3.2.1-amzn-6	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.2.1-amzn-6	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-6	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.4-amzn-2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.4-amzn-2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.4-amzn-2	HBase client en ligne de commande.
hbase-rest-server	2.4.4-amzn-2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.4-amzn-2	Service fournissant un point de terminaison Thrift pour. HBase
hbase-operator-tools	2.4.4-amzn-2	Outil de réparation pour les HBase clusters Apache.

Composant	Version	Description
hcatalog-client	3.1.2-amzn-7	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.2-amzn-7	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.2-amzn-7	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.2-amzn-7	Client de ligne de commande Hive.
hive-hbase	3.1.2-amzn-7	Client Hive-hbase.
hive-metastore-server	3.1.2-amzn-7	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.2-amzn-7	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.10.1-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.10.1-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.

Composant	Version	Description
hudi-trino	0.10.1-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.10.1-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.10.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
iceberg	0.13.1	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.8.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	11,0,194	Pilotes Nvidia et boîte à outils Cuda

Composant	Version	Description
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.2	Les bibliothèques phoenix pour serveur et client
phoenix-connectors	5.1.2	Connecteurs Apache Phoenix pour Spark-3
phoenix-query-server	5.1.2	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.267-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.267-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.267-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.

Composant	Version	Description
trino-coordinator	367-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	367-amzn-0	Service pour exécuter les éléments d'une requête.
trino-client	367-amzn-0	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.2.0-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.2.0-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.2.0-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.2.0-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.

Composant	Version	Description
spark-rapids	22.02.0-amzn-0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0.10.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.7	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.5.7	ZooKeeper client en ligne de commande.

Classifications des configurations 6.6.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.6.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.

Classifications	Description	Actions de reconfiguration
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-iceberg	Modifiez les valeurs dans le fichier iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.

Classifications	Description	Actions de reconfiguration
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Amazon EMR version 6.5.0

Versions d’application 6.5.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.5.0	emr-6.4.0	emr-6.3.1	emr-6.3.0
AWS SDK pour Java	1,12,31	1,12,31	1,1,977	1,1,977
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,10	2,1,10	2,1,10	2,1,10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.0	1.13.1	1.12.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.2.6	2.2.6
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.9.0-amzn-1	0.8.0-amzn-0	0.7.0-amzn-0	0.7.0-amzn-0

	emr-6.5.0	emr-6.4.0	emr-6.3.1	emr-6.3.0
Hue	4.9.0	4.9.0	4.9.0	4.9.0
Iceberg	0.12.0	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.2.2	1.2.2
Livy	0.7.1	0.7.1	0.7.0	0.7.0
MXNet	1.8.0	1.8.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,261	0,254,1	0,245,1	0,245,1
Spark	3.1.2	3.1.2	3.1.1	3.1.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	360	359	350	350
Zeppelin	0.10.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.4.14	3.4.14

Notes de mise à jour 6.5.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.5.0. Les modifications ont été apportées à la version 6.4.0.

Date de parution initiale : 20 janvier 2022

Date de mise à niveau : 21 mars 2022

Nouvelles fonctions

- [Mise à l'échelle gérée] Optimisation de la mise à l'échelle gérée des données de réorganisation Spark – Pour Amazon EMR versions 5.34.0 et ultérieures, et EMR versions 6.4.0 et ultérieures, la mise à l'échelle gérée prend désormais en compte les données de réorganisation Spark (données que Spark redistribue entre les partitions pour effectuer des opérations spécifiques). Pour plus d'informations sur les opérations de réorganisation, consultez Utilisation de la mise à l'échelle gérée par EMR dans Amazon EMR dans le Guide de gestion Amazon EMR et le Guide de programmation Spark.
- À partir d'Amazon EMR 5.32.0 et 6.5.0, le dimensionnement dynamique de l'exécuteur pour Apache Spark est activé par défaut. Pour activer ou désactiver cette fonctionnalité, vous pouvez utiliser le paramètre de configuration `spark.yarn.heterogeneousExecutors.enabled`.
- Prise en charge du format de table ouvert Apache Iceberg pour les jeux de données analytiques volumineux.
- Support pour ranger-trino-plugin 2.0.1-amzn-1
- Prise en charge de toree 0.5.0

Modifications, améliorations et problèmes résolus

- La version 6.5 d'Amazon EMR prend désormais en charge Apache Iceberg 0.12.0 et apporte des améliorations d'exécution avec l'environnement d'exécution Amazon EMR pour Apache Spark, l'environnement d'exécution Amazon EMR pour Presto et l'environnement d'exécution Amazon EMR pour Apache Hive.
- Apache Iceberg est un format de table ouvert pour les grands jeux de données dans Amazon S3. Il fournit des performances de requête rapides sur de grandes tables, des validations atomiques, des écritures simultanées et une évolution de table compatible avec SQL. Avec EMR 6.5, vous pouvez utiliser Apache Spark 3.1.2 avec le format de table Iceberg.

- Apache Hudi 0.9 ajoute la prise en charge de DDL et DML de Spark SQL. Cela vous permet de créer et de modifier des tables Hudi en utilisant uniquement des instructions SQL. Apache Hudi 0.9 inclut également des améliorations des performances côté requête et côté écriture.
- L'environnement d'exécution Amazon EMR pour Apache Hive améliore les performances d'Apache Hive sur Amazon S3 en supprimant les opérations de changement de nom pendant les opérations intermédiaires et en améliorant les performances des commandes de vérification du métastore (MSCK) utilisées pour la réparation des tables.

Problèmes connus

- Lorsque les versions 6.5.0, 6.6.0 ou 6.7.0 d'Amazon EMR lisent les tables Apache Phoenix via le shell Apache Spark, une erreur `NoSuchMethodError` se produit car Amazon EMR utilise une `Hbase.compat.version` incorrecte. La version 6.8.0 d'Amazon EMR résout ce problème.
- Les clusters de la solution groupée Hbase en haute disponibilité (HA) ne parviennent pas à se provisionner avec la taille de volume et le type d'instance par défaut. La solution à ce problème consiste à augmenter la taille du volume racine.
- Pour utiliser les actions Spark avec Apache Oozie, vous devez ajouter la configuration suivante à votre fichier Oozie `workflow.xml`. Sinon, plusieurs bibliothèques critiques telles que Hadoop et EMRFS seront absentes du classpath des exécuteurs Spark lancés par Oozie.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 6.5.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	3.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.4.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.19.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.1.0	Connecteur S3Select EMR
emrfs	2,48,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.14.0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.14.0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	3.2.1-amzn-5	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-5	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-5	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.2.1-amzn-5	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-5	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.2.1-amzn-5	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-5	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-5	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.2.1-amzn-5	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	3.2.1-amzn-5	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-5	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.4-amzn-1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.4-amzn-1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.4-amzn-1	HBase client en ligne de commande.
hbase-rest-server	2.4.4-amzn-1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.4.4-amzn-1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	3.1.2-amzn-6	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	3.1.2-amzn-6	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.2-amzn-6	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.2-amzn-6	Client de ligne de commande Hive.
hive-hbase	3.1.2-amzn-6	Client Hive-hbase.
hive-metastore-server	3.1.2-amzn-6	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.2-amzn-6	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.9.0-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.9.0-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.9.0-amzn-1	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.

Composant	Version	Description
hudi-spark	0.9.0-amzn-1	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.9.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
iceberg	0.12.0	Apache Iceberg est un format de table ouvert pour les jeux de données analytiques de grande taille.
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.8.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	10,1243	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.2	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	5.1.2	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.261-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.261-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.261-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	360	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	360	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
trino-client	360	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.1.2-amzn-1	Clients de ligne de commande Spark.
spark-history-server	3.1.2-amzn-1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.1.2-amzn-1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.1.2-amzn-1	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	0,4.1	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'application tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0.10.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.7	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.5.7	ZooKeeper client en ligne de commande.

Classifications des configurations 6.5.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.5.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Classifications	Description	Actions de reconfiguration
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
iceberg-defaults	Modification des valeurs du fichier iceberg-defaults.conf d'Iceberg.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.

Classifications	Description	Actions de reconfiguration
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)

Classifications	Description	Actions de reconfiguration
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.

Classifications	Description	Actions de reconfiguration
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.

Classifications	Description	Actions de reconfiguration
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Amazon EMR version 6.4.0

Versions d’application 6.4.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.4.0	emr-6.3.1	emr-6.3.0	emr-6.2.1
AWS SDK pour Java	1,12,31	1,1,977	1,1,977	1,1880
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,10	2,1,10	2,1,10	2,1,10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.13.1	1.12.1	1.12.1	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.2.6	2.2.6	2.2.6-amzn-0
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.8.0-amzn-0	0.7.0-amzn-0	0.7.0-amzn-0	0.6.0-amzn-1

	emr-6.4.0	emr-6.3.1	emr-6.3.0	emr-6.2.1
Hue	4.9.0	4.9.0	4.9.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.2.2	1.2.2	1.1.0
Livy	0.7.1	0.7.0	0.7.0	0.7.0
MXNet	1.8.0	1.7.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.0
Phoenix	5.1.2	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,254,1	0,245,1	0,245,1	0,238,3
Spark	3.1.2	3.1.1	3.1.1	3.0.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	359	350	350	343
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.4.14	3.4.14	3.4.14

Notes de mise à jour 6.4.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.4.0. Les modifications ont été apportées à la version 6.3.0.

Date de parution initiale : 20 septembre 2021

Date de mise à niveau : 21 mars 2022

Applications prises en charge

- AWS SDK pour Java version 1.12.31
- CloudWatch Version 2.2.0 de l'évier
- Connecteur DynamoDB version 4.16.0
- EMRFS version 2.47.0
- Amazon EMR Goodies version 3.2.0
- Amazon EMR Kinesis Connector version 3.5.0
- Amazon EMR Record Server version 2.1.0
- Amazon EMR Scripts version 2.5.0
- Flink version 1.13.1
- Ganglia version 3.7.2
- AWS Client Glue Hive Metastore version 3.3.0
- Hadoop version 3.2.1-amzn-4
- HBase version 2.4.4-amzn-0
- HBase-operator-tools 1.1.0
- HCatalog version 3.1.2-amzn-5
- Hive version 3.1.2-amzn-5
- Hudi version 0.8.0-amzn-0
- Hue version 4.9.0
- Java JDK version Corretto-8.302.08.1 (build 1.8.0_302-b08)
- JupyterHub version 1.4.1
- Livy version 0.7.1-incubating
- MXNet version 1.8.0
- Oozie version 5.2.1

- Phoenix version 5.1.2
- Pig version 0.17.0
- Presto version 0.254.1-amzn-0
- Trino version 359
- Apache Ranger KMS (chiffrement transparent à plusieurs maîtres) version 2.0.0
- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.2.0
- SageMaker Version 1.4.1 du SDK Spark
- Scala version 2.12.10 (machine virtuelle du serveur OpenJDK 64 bits, Java 1.8.0_282)
- Spark version 3.1.2-amzn-0
- spark-rapids 0.4.1
- Sqoop version 1.4.7
- TensorFlow version 2.4.1
- tez version 0.9.2
- Zeppelin version 0.9.0
- Zookeeper version 3.5.7
- Connecteurs et pilotes : Connecteur DynamoDB 4.16.0

Nouvelles fonctionnalités

- [Mise à l'échelle gérée] Optimisation de la mise à l'échelle gérée des données de réorganisation Spark – Pour Amazon EMR versions 5.34.0 et ultérieures, et EMR versions 6.4.0 et ultérieures, la mise à l'échelle gérée prend désormais en compte les données de réorganisation Spark (données que Spark redistribue entre les partitions pour effectuer des opérations spécifiques). Pour plus d'informations sur les opérations de réorganisation, consultez Utilisation de la mise à l'échelle gérée par EMR dans Amazon EMR dans le Guide de gestion Amazon EMR et le Guide de programmation Spark.
- Sur les clusters Amazon EMR compatibles avec Apache Ranger, vous pouvez utiliser Apache Spark SQL pour insérer des données dans les tables de métastore Apache Hive ou les mettre à jour à l'aide de `INSERT INTO`, `INSERT OVERWRITE` et `ALTER TABLE`. Lorsque vous utilisez ALTER TABLE avec Spark SQL, l'emplacement d'une partition doit être le répertoire enfant d'un emplacement de table. Amazon EMR ne prend actuellement pas en charge l'insertion de données dans une partition où l'emplacement de la partition est différent de celui de la table.

- PrestoSQL a été renommé Trino.
- Hive : l'exécution de requêtes SELECT simples avec la clause LIMIT est accélérée en arrêtant l'exécution de la requête dès que le nombre d'enregistrements mentionné dans la clause LIMIT est récupéré. Les requêtes SELECT simples sont des requêtes qui ne contiennent pas de clause GROUP BY/ORDER BY ou des requêtes qui n'ont pas d'étape de réduction. Par exemple, `SELECT * from <TABLE> WHERE <Condition> LIMIT <Number>`.

Contrôles de simultanéité Hudi

- Hudi prend désormais en charge le contrôle de simultanéité optimiste (OCC), qui peut être exploité avec des opérations d'écriture telles que UPSERT et INSERT pour permettre les modifications de plusieurs enregistreurs sur la même table Hudi. Il s'agit d'un OCC au niveau du fichier, de sorte que deux validations (ou enregistreurs) peuvent écrire dans la même table, si leurs modifications n'entrent pas en conflit. Pour plus d'informations, consultez le contrôle de simultanéité de Hudi.
- Zookeeper est installé sur les clusters Amazon EMR, qui peut être utilisé comme fournisseur de verrous pour OCC. Pour faciliter l'utilisation de cette fonctionnalité, les propriétés préconfigurées des clusters Amazon EMR sont les suivantes :

```
hoodie.write.lock.provider=org.apache.hudi.client.transaction.lock.ZookeeperBasedLockProvider
hoodie.write.lock.zookeeper.url=<EMR Zookeeper URL>
hoodie.write.lock.zookeeper.port=<EMR Zookeeper Port>
hoodie.write.lock.zookeeper.base_path=/hudi
```

  Pour activer l'OCC, vous devez configurer les propriétés suivantes soit avec leurs options de tâche Hudi, soit au niveau du cluster à l'aide de l'API de configuration Amazon EMR :

```
hoodie.write.concurrency.mode=optimistic_concurrency_control
hoodie.cleaner.policy.failed.writes=LAZY (Performs cleaning of failed writes lazily
 instead of inline with every write)
hoodie.write.lock.zookeeper.lock_key=<Key to uniquely identify the Hudi table> (Table
 Name is a good option)
```

Hudi Monitoring : CloudWatch intégration d'Amazon pour générer des rapports sur Hudi Metrics

- Amazon EMR prend en charge la publication de Hudi Metrics sur Amazon. CloudWatch Elle est activée en définissant les configurations requises suivantes :

```
hoodie.metrics.on=true
hoodie.metrics.reporter.type=CLOUDWATCH
```

- Les configurations Hudi facultatives que vous pouvez modifier sont les suivantes :

Paramètre	Description	Valeur
hoodie.metrics.cloudwatch.report.period.seconds	Fréquence (en secondes) à laquelle les statistiques doivent être communiquées à Amazon CloudWatch	La valeur par défaut est de 60 s, ce qui convient à la résolution d'une minute par défaut proposée par Amazon CloudWatch
hoodie.metrics.cloudwatch.metric.prefix	Préfixe à ajouter au nom de chaque métrique	La valeur par défaut est vide (pas de préfixe)
hoodie.metrics.cloudwatch.namespace	CloudWatch Espace de noms Amazon sous lequel les métriques sont publiées	La valeur par défaut est Hudi
hoodie.metrics.cloudwatch.maxDatumsPerDemande	Nombre maximum de références à inclure dans une demande adressée à Amazon CloudWatch	La valeur par défaut est 20, identique à la valeur CloudWatch par défaut d'Amazon

Prise en charge et amélioration des configurations Hudi d'Amazon EMR

- Les clients peuvent désormais tirer parti de l'API de configuration EMR et de la fonctionnalité de reconfiguration pour configurer les configurations Hudi au niveau du cluster. Un nouveau support de configuration basé sur des fichiers a été introduit via/etc/hudi/conf/hudi-defaults.conf, à l'instar d'autres applications telles que Spark, Hive, etc. EMR configure quelques paramètres par défaut pour améliorer l'expérience utilisateur :

  — `hoodie.datasource.hive_sync.jdbcurl` est configuré selon l'URL du serveur Hive du cluster et n'a plus besoin d'être spécifié. Cela est particulièrement utile lorsque vous exécutez une tâche en mode cluster Spark, où vous deviez auparavant spécifier l'adresse IP principale Amazon EMR.

— HBase des configurations spécifiques, utiles pour utiliser l' HBase index avec Hudi.

— Configuration spécifique au fournisseur de verrous Zookeeper, comme indiqué dans la section Contrôle de simultanéité, qui facilite l'utilisation du contrôle de simultanéité optimiste (OCC).

- Des modifications supplémentaires ont été introduites pour réduire le nombre de configurations à transmettre et pour en déduire automatiquement dans la mesure du possible :

  — Le mot clé `partitionBy` peut être utilisé pour spécifier la colonne de partition.

  — Lorsque vous activez Hive Sync, il n'est plus obligatoire de spécifier `HIVE_TABLE_OPT_KEY, HIVE_PARTITION_FIELDS_OPT_KEY, HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY`. Ces valeurs peuvent être déduites du nom de la table Hudi et du champ de partition.

  — `KEYGENERATOR_CLASS_OPT_KEY` n'est pas obligatoire et peut être déduite de cas plus simples de `SimpleKeyGenerator` et`ComplexKeyGenerator`.

Mises en garde de Hudi

- Hudi ne prend pas en charge l'exécution vectorisée dans les tables Hive for Merge on Read (MoR) et Bootstrap. Par exemple, `count(*)` échoue avec la table en temps réel de Hudi lorsque `hive.vectorized.execution.enabled` est défini sur true. Comme solution de contournement, vous pouvez désactiver la lecture vectorisée en définissant `hive.vectorized.execution.enabled` sur `false`.
- La prise en charge des enregistreurs multiples n'est pas compatible avec la fonction d'amorçage de Hudi.
- Flink Streamer et Flink SQL sont des fonctionnalités expérimentales dans cette version. Ces fonctionnalités ne sont pas recommandées pour les déploiements de production.

Modifications, améliorations et problèmes résolus

Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.

- Auparavant, le redémarrage manuel du gestionnaire de ressources sur un cluster multimaître provoquait le rechargement par les démons Amazon EMR on-cluster, comme Zookeeper, de tous les nœuds précédemment mis hors service ou perdus dans le fichier znode de Zookeeper. Cela a entraîné le dépassement des limites par défaut dans certaines situations. Amazon EMR supprime

désormais les enregistrements de nœuds mis hors service ou perdus datant de plus d'une heure du fichier Zookeeper et les limites internes ont été augmentées.

- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Configuration d'un cluster pour résoudre les problèmes de performances d'Apache YARN Timeline Server versions 1 et 1.5

  Les versions 1 et 1.5 d'Apache YARN Timeline Server peuvent entraîner des problèmes de performances avec de grands clusters EMR très actifs, en particulier avec `yarn.resourcemanager.system-metrics-publisher.enabled=true`, le paramètre par défaut d'Amazon EMR. YARN Timeline Server v2 open source résout le problème de performance lié à la capacité de mise à l'échelle de YARN Timeline Server.

  Les autres solutions à ce problème incluent :

  - Configuration de yarn.resourcemanager. system-metrics-publisher.enabled=false dans le fichier yarn-site.xml.

- Activation du correctif pour ce problème lors de la création d'un cluster, comme décrit ci-dessous.

  Les versions Amazon EMR suivantes contiennent un correctif pour ce problème de performance de YARN Timeline Server.

  EMR 5.30.2, 5.31.1, 5.32.1, 5.33.1, 5.34.x, 6.0.1, 6.1.1, 6.2.1, 6.3.1, 6.4.x

  Pour activer le correctif sur l'une des versions Amazon EMR spécifiées ci-dessus, définissez ces propriétés sur `true` dans un fichier JSON de configuration transmis à l'aide du paramètre de commande `aws emr create-cluster` : `--configurations file://./configurations.json`. Vous pouvez également activer le correctif à l'aide de l'interface utilisateur de la console de reconfiguration.

  Exemple du contenu du fichier configurations.json :

```
[
{
"Classification": "yarn-site",
"Properties": {
"yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
 "true",
"yarn.resourcemanager.system-metrics-publisher.enabled": "true"
},
"Configurations": []
}
]
```

- WebHDFS et le serveur HttpFS sont désactivés par défaut. Vous pouvez réactiver WebHDFS en utilisant la configuration Hadoop, `dfs.webhdfs.enabled`. Le serveur HttpFS peut être démarré en utilisant `sudo systemctl start hadoop-httpfs`.
- Le protocole HTTPS est désormais activé par défaut pour les référentiels Amazon Linux. Si vous utilisez une politique VPCE Amazon S3 pour restreindre l'accès à des compartiments spécifiques, vous devez ajouter le nouvel ARN du compartiment Amazon Linux `arn:aws:s3:::amazonlinux-2-repos-$region/*` à votre politique (remplacez `$region` par la région où se trouve le point de terminaison). Pour plus d'informations, consultez cette rubrique dans les forums de AWS discussion. Annonce : Amazon Linux 2 permet désormais d'utiliser le protocole HTTPS lors de la connexion aux référentiels de packages.

- Hive : les performances des requêtes d'écriture sont améliorées en permettant l'utilisation d'un répertoire temporaire sur HDFS pour la dernière tâche. Les données temporaires pour la tâche finale sont écrites sur HDFS au lieu d'Amazon S3 et les performances sont améliorées car les données sont déplacées de HDFS vers l'emplacement de la table finale (Amazon S3) au lieu d'être déplacées entre les appareils Amazon S3.
- Hive : amélioration du temps de compilation des requêtes jusqu'à 2,5 fois avec l'élimination des partitions du métastore Glue.
- Par défaut, lorsque les éléments intégrés UDFs sont transmis par Hive au serveur Hive Metastore, seul un sous-ensemble de ces éléments intégrés UDFs est transmis au Glue Metastore, car Glue ne prend en charge que des opérateurs d'expression limités. Si vous définissez `hive.glue.partition.pruning.client=true`, tout l'élimination des partitions se fait du côté client. Si vous définissez `hive.glue.partition.pruning.server=true`, tout l'élimination des partitions se fait du côté serveur.

Problèmes connus

- Les requêtes Hue ne fonctionnent pas dans Amazon EMR 6.4.0 car le serveur Apache Hadoop HTTPFS est désactivé par défaut. Pour utiliser Hue sur Amazon EMR 6.4.0, démarrez manuellement le serveur HttpFS sur le nœud primaire d'Amazon EMR à l'aide de `sudo systemctl start hadoop-httpfs`, ou utilisez une étape d'Amazon EMR.
- La fonctionnalité Blocs-notes Amazon EMR utilisée avec l'emprunt d'identité de l'utilisateur Livy ne fonctionne pas car HttpFS est désactivé par défaut. Dans ce cas, le bloc-notes EMR ne peut pas se connecter au cluster dont l'emprunt d'identité Livy est activé. La solution consiste à démarrer le serveur HttpFS avant de connecter le bloc-notes EMR au cluster à l'aide de `sudo systemctl start hadoop-httpfs`.
- Dans la version 6.4.0 d'Amazon EMR, Phoenix ne prend pas en charge le composant des connecteurs Phoenix.
- Pour utiliser les actions Spark avec Apache Oozie, vous devez ajouter la configuration suivante à votre fichier Oozie `workflow.xml`. Sinon, plusieurs bibliothèques critiques telles que Hadoop et EMRFS seront absentes du classpath des exécuteurs Spark lancés par Oozie.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster

de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :

- Deux partitions ou plus sont analysées à partir de la même table.
- Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
- Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 6.4.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark

Composant	Version	Description
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.2.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.3.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.18.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.1.0	Connecteur S3Select EMR
emrfs	2,47,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.13.1	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.13.1	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.2.1-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.2.1-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-4	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.2.1-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	3.2.1-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.2.1-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.2.1-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.4.4-amzn-0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.4.4-amzn-0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.4.4-amzn-0	HBase client en ligne de commande.
hbase-rest-server	2.4.4-amzn-0	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	2.4.4-amzn-0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	3.1.2-amzn-5	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.2-amzn-5	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.2-amzn-5	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.2-amzn-5	Client de ligne de commande Hive.
hive-hbase	3.1.2-amzn-5	Client Hive-hbase.
hive-metastore-server	3.1.2-amzn-5	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.2-amzn-5	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.8.0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.

Composant	Version	Description
hudi-presto	0.8.0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-trino	0.8.0-amzn-0	Bibliothèque de solution groupée pour exécuter Trino avec Hudi.
hudi-spark	0.8.0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.9.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.8.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	10,1243	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5.1.2	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	5.1.2	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.254.1-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.254.1-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.254.1-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
trino-coordinator	359	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants trino-workers.
trino-worker	359	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
trino-client	359	Client de ligne de commande Trino installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Trino n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.1.2-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.1.2-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.1.2-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.1.2-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	0,4.1	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'application tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0.9.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.5.7	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.5.7	ZooKeeper client en ligne de commande.

Classifications des configurations 6.4.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.4.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Classifications	Description	Actions de reconfiguration
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.

Classifications	Description	Actions de reconfiguration
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Classifications	Description	Actions de reconfiguration
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
trino-log	Modifiez les valeurs dans le fichier log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Modifiez les valeurs dans le fichier config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Modification des valeurs dans le fichier trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Modifiez les valeurs dans le fichier node.properties de Trino.	Not available.
trino-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Trino.	Not available.
trino-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Trino.	Not available.
trino-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Trino.	Not available.
trino-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Trino.	Not available.
trino-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Trino.	Not available.
trino-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Trino.	Not available.
trino-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Trino.	Not available.
trino-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Trino.	Not available.
trino-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Trino.	Not available.
trino-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Trino.	Not available.

Classifications	Description	Actions de reconfiguration
trino-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Trino.	Not available.
trino-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.

Classifications	Description	Actions de reconfiguration
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

Amazon EMR version 6.3.1

Versions d’application 6.3.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, PrestoSQL, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR

- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.3.1	emr-6.3.0	emr-6.2.1	emr-6.2.0
AWS SDK pour Java	1,1,977	1,1,977	1,1880	1,1880
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,10	2,1,10	2,1,10	2,1,10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.12.1	1.11.2	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6	2.2.6	2.2.6-amzn-0	2.2.6-amzn-0
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.7.0-amzn-0	0.7.0-amzn-0	0.6.0-amzn-1	0.6.0-amzn-1
Hue	4.9.0	4.9.0	4.8.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0

	emr-6.3.1	emr-6.3.0	emr-6.2.1	emr-6.2.0
JupyterHub	1.2.2	1.2.2	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,245,1	0,245,1	0,238,3	0,238,3
Spark	3.1.1	3.1.1	3.0.1	3.0.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.3.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	350	350	343	343
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 6.3.1

Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.

Modifications, améliorations et problèmes résolus

- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Le protocole HTTPS est désormais activé par défaut pour les référentiels Amazon Linux. Si vous utilisez une politique VPCE Amazon S3 pour restreindre l'accès à des compartiments spécifiques, vous devez ajouter le nouvel ARN du compartiment Amazon Linux `arn:aws:s3:::amazonlinux-2-repos-$region/*` à votre politique (remplacez `$region` par la région où se trouve le point de terminaison). Pour plus d'informations, consultez cette rubrique dans les forums de AWS discussion. Annonce : Amazon Linux 2 permet désormais d'utiliser le protocole HTTPS lors de la connexion aux référentiels de packages.

Problèmes connus

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster

de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :

- Deux partitions ou plus sont analysées à partir de la même table.
- Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
- Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 6.3.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark

Composant	Version	Description
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.2.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.2.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.18.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.1.0	Connecteur S3Select EMR
emrfs	2,46,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.12.1	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.12.1	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.2.1-amzn-3.1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-3.1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-3.1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.2.1-amzn-3.1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-3.1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.2.1-amzn-3.1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-3.1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	3.2.1-amzn-3.1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.2.1-amzn-3.1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.2.1-amzn-3.1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-3.1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.2.6-amzn-1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.2.6-amzn-1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.2.6-amzn-1	HBase client en ligne de commande.
hbase-rest-server	2.2.6-amzn-1	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	2.2.6-amzn-1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	3.1.2-amzn-4	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.2-amzn-4	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.2-amzn-4	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.2-amzn-4	Client de ligne de commande Hive.
hive-hbase	3.1.2-amzn-4	Client Hive-hbase.
hive-metastore-server	3.1.2-amzn-4	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.2-amzn-4	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.7.0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.

Composant	Version	Description
hudi-presto	0.7.0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-prestosql	0.7.0-amzn-0	Bibliothèque de solution groupée pour exécuter PrestoSQL avec Hudi.
hudi-spark	0.7.0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.9.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.2.2	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.7.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.
nvidia-cuda	10,1243	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5,0,0- -2,0 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	5,0,0- -2,0 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.245.1-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.245.1-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.245.1-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
prestosql-coordinator	350	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants prestosql-workers.
prestosql-worker	350	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
prestosql-client	350	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.1.1-amzn-0.1	Clients de ligne de commande Spark.
spark-history-server	3.1.1-amzn-0.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.1.1-amzn-0.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.1.1-amzn-0.1	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	0,4.1	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'application tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0.9.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 6.3.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.3.1

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Classifications	Description	Actions de reconfiguration
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.

Classifications	Description	Actions de reconfiguration
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
prestosql-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoSQL)

Classifications	Description	Actions de reconfiguration
prestosql-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	Modifiez les valeurs dans le fichier node.properties de PrestoSQL.	Not available.
prestosql-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de PrestoSQL.	Not available.
prestosql-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de PrestoSQL.	Not available.
prestosql-connector-hive	Modifiez les valeurs dans le fichier hive.properties de PrestoSQL.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de PrestoSQL.	Not available.
prestosql-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de PrestoSQL.	Not available.

Classifications	Description	Actions de reconfiguration
prestosql-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de PrestoSQL.	Not available.
prestosql-connector-memory	Modifiez les valeurs dans le fichier memory.properties de PrestoSQL.	Not available.
prestosql-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de PrestoSQL.	Not available.
prestosql-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de PrestoSQL.	Not available.
prestosql-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de PrestoSQL.	Not available.
prestosql-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de PrestoSQL.	Not available.
prestosql-connector-redis	Modifiez les valeurs dans le fichier redis.properties de PrestoSQL.	Not available.
prestosql-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de PrestoSQL.	Not available.
prestosql-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de PrestoSQL.	Not available.

Classifications	Description	Actions de reconfiguration
prestosql-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de PrestoSQL.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.

Classifications	Description	Actions de reconfiguration
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Amazon EMR version 6.3.0

Versions d’application 6.3.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, PrestoSQL, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.3.0	emr-6.2.1	emr-6.2.0	emr-6.1.1
AWS SDK pour Java	1,1,977	1,1880	1,1880	1,1,828
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,10	2,1,10	2,1,10	2,1,10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.11.2	1.11.2	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6	2.2.6-amzn-0	2.2.6-amzn-0	2.2.5
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.7.0-amzn-0	0.6.0-amzn-1	0.6.0-amzn-1	0.5.2-incubating-amzn-2
Hue	4.9.0	4.8.0	4.8.0	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	-
JupyterHub	1.2.2	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0

	emr-6.3.0	emr-6.2.1	emr-6.2.0	emr-6.1.1
MXNet	1.7.0	1.7.0	1.7.0	1.6.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.0	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,245,1	0,238,3	0,238,3	0.232
Spark	3.1.1	3.0.1	3.0.1	3.0.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.3.1	2.3.1	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	350	343	343	338
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 6.3.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.3.0. Les modifications ont été apportées à la version 6.2.0.

Date de parution initiale : 12 mai 2021

Dernière mise à jour : 9 août 2021

Applications prises en charge

- AWS SDK pour Java version 1.11.977

- CloudWatch Version 2.1.0 de l'évier
- Connecteur DynamoDB version 4.16.0
- EMRFS version 2.46.0
- Amazon EMR Goodies version 3.2.0
- Amazon EMR Kinesis Connector version 3.5.0
- Amazon EMR Record Server version 2.0.0
- Amazon EMR Scripts version 2.5.0
- Flink version 1.12.1
- Ganglia version 3.7.2
- AWS Client Glue Hive Metastore version 3.2.0
- Hadoop version 3.2.1-amzn-3
- HBase version 2.2.6-amzn-1
- HBase-operator-tools 1.0.0
- HCatalog version 3.1.2-amzn-0
- Hive version 3.1.2-amzn-4
- Hudi version 0.7.0-amzn-0
- Hue version 4.9.0
- Java JDK version Corretto-8.282.08.1 (build 1.8.0_282-b08)
- JupyterHub version 1.2.0
- Livy version 0.7.0-incubating
- MXNet version 1.7.0
- Oozie version 5.2.1
- Phoenix version 5.0.0
- Pig version 0.17.0
- Presto version 0.245.1-amzn-0
- PrestoSQL version 350
- Apache Ranger KMS (chiffrement transparent à plusieurs maîtres) version 2.0.0
- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.1.0

- SageMaker Version 1.4.1 du SDK Spark
- Scala version 2.12.10 (machine virtuelle du serveur OpenJDK 64 bits, Java 1.8.0_282)
- Spark version 3.1.1-amzn-0
- spark-rapids 0.4.1
- Sqoop version 1.4.7
- TensorFlow version 2.4.1
- tez version 0.9.2
- Zeppelin version 0.9.0
- Zookeeper version 3.4.14
- Connecteurs et pilotes : Connecteur DynamoDB 4.16.0

Nouvelles fonctionnalités

- Amazon EMR prend en charge les points d'accès Amazon S3, une fonctionnalité d'Amazon S3 qui vous permet de gérer facilement l'accès aux lacs de données partagés. En utilisant votre alias de point d'accès Amazon S3, vous pouvez simplifier l'accès aux données à grande échelle sur Amazon EMR. Vous pouvez utiliser les points d'accès Amazon S3 avec toutes les versions d'Amazon EMR sans frais supplémentaires dans toutes les AWS régions où Amazon EMR est disponible. Pour en savoir plus sur les points d'accès Amazon S3 et les alias de points d'accès, consultez Utilisation d'un alias de type compartiment pour votre point d'accès dans le Guide de l'utilisateur Amazon S3.
- Les nouveaux paramètres d'API `DescribeReleaseLabel` et `ListReleaseLabel` fournissent des détails sur les étiquettes de version d'Amazon EMR. Vous pouvez répertorier par programmation les versions disponibles dans la région où la demande d'API est exécutée, et répertorier les applications disponibles pour une étiquette de version Amazon EMR spécifique. Les paramètres d'étiquette de version répertorient également les versions Amazon EMR qui prennent en charge une application spécifiée, telle que Spark. Ces informations peuvent être utilisées pour lancer par programmation des clusters Amazon EMR. Par exemple, vous pouvez lancer un cluster à l'aide de la dernière version des résultats de `ListReleaseLabel`. Pour plus d'informations, consultez DescribeReleaseLabelet consultez ListReleaseLabelsle manuel Amazon EMR API Reference.
- Avec Amazon EMR 6.3.0, vous pouvez lancer un cluster qui s'intègre nativement à Apache Ranger. Apache Ranger est un cadre open source permettant d'activer, de surveiller et de gérer la sécurité globale des données sur la plateforme Hadoop. Pour plus d'informations, consultez

Apache Ranger. Grâce à l'intégration native, vous pouvez utiliser votre propre Apache Ranger pour appliquer un contrôle précis de l'accès aux données sur Amazon EMR. Consultez Intégration d'Amazon EMR avec Apache Ranger dans le Guide de gestion Amazon EMR.

- Politiques gérées délimitées : pour s'aligner sur les AWS meilleures pratiques, Amazon EMR a introduit des politiques gérées par défaut définies dans la version 2 EMR en remplacement des politiques qui seront déconseillées. Consultez Politiques gérées par Amazon EMR.
- État de prise en charge du service de métadonnées d'instance (IMDS) V2 : pour Amazon EMR 6.2 ou version ultérieure, les composants Amazon EMR sont IMDSv2 utilisés pour tous les appels IMDS. Pour les appels IMDS dans le code de votre application, vous pouvez utiliser les deux IMDSv1 ou configurer l'IMDS pour qu'il ne soit utilisé que IMDSv2 pour renforcer la sécurité. IMDSv2 Si vous la désactivez IMDSv1 dans les versions antérieures d'Amazon EMR 6.x, cela entraîne un échec du démarrage du cluster.

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.
- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas

renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.

- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Le mode explain par défaut de l'interface utilisateur Spark SQL est passé de `extended` à `formatted` dans Spark 3.1. Amazon EMR l'a ramené à `extended` pour inclure des informations sur le plan logique dans l'interface utilisateur Spark SQL. Ceci peut être inversé en définissant `spark.sql.ui.explainMode` sur `formatted`.
- Les commits suivants ont été rétroportés depuis la branche master de Spark.

  - [SPARK-34752][BUILD] Passer Jetty à la version 9.4.37 pour résoudre la CVE-2020-27223.

  - [SPARK-34534] Corrige l'ordre des blocs lorsque vous l'utilisez pour récupérer des blocs. FetchShuffleBlocks

  - [SPARK-34681] [SQL] Correction d'un bogue pour les jointures de hachage réorganisées externes complètes lorsque le côté gauche est créé avec une condition non égale.

  - [SPARK-34497] [SQL] Correction des fournisseurs de connexion JDBC intégrés pour restaurer les changements de contexte de sécurité de la JVM.
- Pour améliorer l'interopérabilité avec le RAPIDs plugin Nvidia Spark, une solution de contournement a été ajoutée pour résoudre un problème empêchant le déclenchement de l'élagage dynamique des partitions lors de l'utilisation de Nvidia Spark RAPIDs alors que l'exécution adaptative des requêtes est désactivée, consultez les numéros #1378 de RAPIDS et #1386 de RAPIDS. Pour plus de détails sur la nouvelle configuration `spark.sql.optimizer.dynamicPartitionPruning.enforceBroadcastReuse`, consultez Problème RAPIDS n°1386.
- L'algorithme par défaut du validateur de sortie de fichier a été modifié de l'algorithme v2 à l'algorithme v1 dans Spark 3.1 open source. Pour plus d'informations, consultez cette page Amazon EMR optimisant les performances de Spark – élimination dynamique des partitions.
- Amazon EMR est revenu à l'algorithme v2, l'algorithme par défaut utilisé dans les versions antérieures d'Amazon EMR 6.x, pour éviter une régression des performances. Pour rétablir le comportement de Spark 3.1 open source, définissez `spark.hadoop.mapreduce.fileoutputcommitter.algorithm.version` sur 1. Le Spark

open source a apporté cette modification parce que la validation des tâches dans l'algorithme v2 du validateur de sortie de fichier n'est pas atomique, ce qui peut entraîner un problème d'exactitude des données de sortie dans certains cas. Cependant, la validation des tâches dans l'algorithme v1 n'est pas non plus atomique. Dans certains scénarios, la validation des tâches comprend une suppression effectuée avant un changement de nom. Cela peut entraîner un problème silencieux d'exactitude des données.

- Correction des problèmes de Managed Scaling dans les versions précédentes d'Amazon EMR et améliorations permettant de réduire considérablement les taux d'échec des applications.
- Le bundle AWS Java SDK a été installé sur chaque nouveau cluster. Il s'agit d'un fichier jar unique contenant tous les services SDKs et leurs dépendances, au lieu de fichiers jar de composants individuels. Pour plus d'informations, consultez Dépendance du kit SDK Java groupé.

Problèmes connus

- Pour les clusters de sous-réseaux privés Amazon EMR 6.3.0 et 6.2.0, vous ne pouvez pas accéder à l'interface utilisateur Web de Ganglia. Vous recevrez un message d'erreur « accès refusé (403) ». D'autres sites Web UIs, tels que Spark, Hue JupyterHub, Zeppelin, Livy et Tez, fonctionnent normalement. L'accès à l'interface utilisateur Web de Ganglia sur les clusters de sous-réseaux publics fonctionne également normalement. Pour résoudre ce problème, redémarrez le service httpd sur le nœud primaire avec `sudo systemctl restart httpd`. Ce problème est résolu dans Amazon EMR 6.4.0.
- Lorsque AWS Glue Data Catalog est activé, l'utilisation de Spark pour accéder à une base de données AWS Glue avec un URI d'emplacement de chaîne nul peut échouer. Cela se produit pour les versions antérieures d'Amazon EMR, mais pour SPARK-31709 (https://issues.apache). org/jira/browse/SPARK-31709) le rend applicable à un plus grand nombre de cas. Par exemple, lorsque vous créez une table dans la base de données AWS Glue par défaut dont l'URI d'emplacement est une chaîne nulle, le message « Impossible de créer un chemin à partir d'une chaîne vide » `spark.sql("CREATE TABLE mytest (key string) location '/table_path';")` échoue. Pour contourner ce problème, définissez manuellement l'URI de localisation de vos bases de données AWS Glue, puis créez des tables dans ces bases de données à l'aide de Spark.
- Dans Amazon EMR 6.3.0, PrestoSQL a mis à niveau de la version 343 à la version 350. Il y a deux changements liés à la sécurité de l'open source qui se rapportent à ce changement de version. Le contrôle d'accès au catalogue basé sur les fichiers passe de `deny` à `allow` lorsque les règles de propriété de table, de schéma ou de session ne sont pas définies. De même, le contrôle d'accès au système basé sur les fichiers est modifié pour prendre en charge les fichiers sans règles de catalogue définies. Dans ce cas, tous les accès aux catalogues sont autorisés.

Pour plus d'informations, consultez Version 344 (9 octobre 2020).

- Notez que le répertoire utilisateur Hadoop (/) home/hadoop) is readable by everyone. It has Unix 755 (drwxr-xr-x) directory permissions to allow read access by frameworks like Hive. You can put files in /home/hadoop et ses sous-répertoires, mais soyez conscient des autorisations associées à ces répertoires afin de protéger les informations sensibles.
- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Redémarrer InstanceController

     ```
     $ sudo systemctl daemon-reload
     ```

```
$ sudo systemctl restart instance-controller
```

Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Les clusters EMR qui exécutent Amazon Linux ou Amazon Linux 2 Amazon Machine Images (AMIs) utilisent le comportement par défaut d'Amazon Linux et ne téléchargent ni n'installent automatiquement les mises à jour importantes et critiques du noyau nécessitant un redémarrage. Ce comportement est identique à celui des autres EC2 instances Amazon qui exécutent l'AMI Amazon Linux par défaut. Si de nouvelles mises à jour logicielles Amazon Linux nécessitant un redémarrage (telles que les mises à jour du noyau, de NVIDIA et de CUDA) sont disponibles après la publication d'une version d'Amazon EMR, les instances de cluster EMR qui exécutent l'AMI par défaut ne téléchargent pas et n'installent pas automatiquement ces mises à jour. Pour obtenir les mises à jour du noyau,

vous pouvez personnaliser votre AMI Amazon EMR afin d'utiliser la dernière AMI Amazon Linux.

- Pour utiliser les actions Spark avec Apache Oozie, vous devez ajouter la configuration suivante à votre fichier Oozie `workflow.xml`. Sinon, plusieurs bibliothèques critiques telles que Hadoop et EMRFS seront absentes du classpath des exécuteurs Spark lancés par Oozie.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 6.3.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.2.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.2.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.18.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.1.0	Connecteur S3Select EMR
emrfs	2,46,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.12.1	Applications et scripts client de la ligne de commande Apache Flink.

Composant	Version	Description
flink-jobmanager-config	1.12.1	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.2.1-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.2.1-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-3	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.

Composant	Version	Description
hadoop-httpfs-server	3.2.1-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.2.1-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.2.1-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.2.6-amzn-1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.2.6-amzn-1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.2.6-amzn-1	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	2.2.6-amzn-1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.2.6-amzn-1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	3.1.2-amzn-4	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.2-amzn-4	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.2-amzn-4	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.2-amzn-4	Client de ligne de commande Hive.
hive-hbase	3.1.2-amzn-4	Client Hive-hbase.
hive-metastore-server	3.1.2-amzn-4	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.2-amzn-4	Service pour l'acceptation de requêtes Hive en tant que requêtes web.

Composant	Version	Description
hudi	0.7.0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.7.0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-prestosql	0.7.0-amzn-0	Bibliothèque de solution groupée pour exécuter PrestoSQL avec Hudi.
hudi-spark	0.7.0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.9.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.2.2	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.7.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MariaDB.

Composant	Version	Description
nvidia-cuda	10,1243	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5,0,0- -2,0 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	5,0,0- -2,0 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.245.1-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.245.1-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.245.1-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.

Composant	Version	Description
prestosql-coordinator	350	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants prestosql-workers.
prestosql-worker	350	Service pour exécuter les éléments d'une requête.
prestosql-client	350	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.1.1-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.1.1-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.1.1-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.1.1-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.

Composant	Version	Description
spark-rapids	0,4.1	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0.9.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 6.3.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.3.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Restarts Flink history server.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.

Classifications	Description	Actions de reconfiguration
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
prestosql-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoSQL)

Classifications	Description	Actions de reconfiguration
prestosql-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	Modifiez les valeurs dans le fichier node.properties de PrestoSQL.	Not available.
prestosql-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de PrestoSQL.	Not available.
prestosql-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de PrestoSQL.	Not available.
prestosql-connector-hive	Modifiez les valeurs dans le fichier hive.properties de PrestoSQL.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de PrestoSQL.	Not available.
prestosql-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de PrestoSQL.	Not available.

Classifications	Description	Actions de reconfiguration
prestosql-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de PrestoSQL.	Not available.
prestosql-connector-memory	Modifiez les valeurs dans le fichier memory.properties de PrestoSQL.	Not available.
prestosql-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de PrestoSQL.	Not available.
prestosql-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de PrestoSQL.	Not available.
prestosql-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de PrestoSQL.	Not available.
prestosql-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de PrestoSQL.	Not available.
prestosql-connector-redis	Modifiez les valeurs dans le fichier redis.properties de PrestoSQL.	Not available.
prestosql-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de PrestoSQL.	Not available.
prestosql-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de PrestoSQL.	Not available.

Classifications	Description	Actions de reconfiguration
prestosql-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de PrestoSQL.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.

Classifications	Description	Actions de reconfiguration
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Amazon EMR version 6.2.1

Versions d’application 6.2.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, PrestoSQL, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.2.1	emr-6.2.0	emr-6.1.1	emr-6.1.0
AWS SDK pour Java	1,1880	1,1880	1,1,828	1,1,828
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,10	2,1,10	2,1,10	2,1,10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.2	1.11.0	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6-amzn-0	2.2.6-amzn-0	2.2.5	2.2.5
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.6.0-amzn-1	0.6.0-amzn-1	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2
Hue	4.8.0	4.8.0	4.7.1	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0

	emr-6.2.1	emr-6.2.0	emr-6.1.1	emr-6.1.0
MXNet	1.7.0	1.7.0	1.6.0	1.6.0
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,238,3	0,238,3	0.232	0.232
Spark	3.0.1	3.0.1	3.0.0	3.0.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.3.1	2.1.0	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	343	343	338	338
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 6.2.1

Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.

Modifications, améliorations et problèmes résolus

- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.

- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Le protocole HTTPS est désormais activé par défaut pour les référentiels Amazon Linux. Si vous utilisez une politique VPCE Amazon S3 pour restreindre l'accès à des compartiments spécifiques, vous devez ajouter le nouvel ARN du compartiment Amazon Linux `arn:aws:s3:::amazonlinux-2-repos-$region/*` à votre politique (remplacez `$region` par la région où se trouve le point de terminaison). Pour plus d'informations, consultez cette rubrique dans les forums de AWS discussion. Annonce : Amazon Linux 2 permet désormais d'utiliser le protocole HTTPS lors de la connexion aux référentiels de packages.

Problèmes connus

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :

- Deux partitions ou plus sont analysées à partir de la même table.
- Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
- Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 6.2.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark

Composant	Version	Description
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.1.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.0.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.16.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.0.0	Connecteur S3Select EMR
emrfs	2,44,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.11.2	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.11.2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.2.1-amzn-2.1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-2.1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-2.1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.2.1-amzn-2.1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-2.1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.2.1-amzn-2.1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-2.1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	3.2.1-amzn-2.1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.2.1-amzn-2.1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.2.1-amzn-2.1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-2.1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.2.6-amzn-0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.2.6-amzn-0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	2.2.6-amzn-0	HBase client en ligne de commande.
hbase-rest-server	2.2.6-amzn-0	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	2.2.6-amzn-0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	3.1.2-amzn-3	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.2-amzn-3	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.2-amzn-3	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.2-amzn-3	Client de ligne de commande Hive.
hive-hbase	3.1.2-amzn-3	Client Hive-hbase.
hive-metastore-server	3.1.2-amzn-3	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.2-amzn-3	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.6.0-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.

Composant	Version	Description
hudi-presto	0.6.0-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-prestosql	0.6.0-amzn-1	Bibliothèque de solution groupée pour exécuter PrestoSQL avec Hudi.
hudi-spark	0.6.0-amzn-1	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.8.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.1.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.7.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5.5.64+	Serveur de base de données MariaDB.
nvidia-cuda	10,1243	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5,0,0- -2,0 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	5,0,0- -2,0 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.238.3-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.238.3-amzn-1	Service pour exécuter les éléments d'une requête.
presto-client	0.238.3-amzn-1	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
prestosql-coordinator	343	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants prestosql-workers.
prestosql-worker	343	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
prestosql-client	343	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.3	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.0.1-amzn-0.1	Clients de ligne de commande Spark.
spark-history-server	3.0.1-amzn-0.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.0.1-amzn-0.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.0.1-amzn-0.1	Bibliothèques Apache Spark requises par les esclaves YARN.
spark-rapids	0.2.0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	2.3.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'application tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0.9.0-preview1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 6.2.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.2.1

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Classifications	Description	Actions de reconfiguration
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Not available.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Not available.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.	Not available.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.

Classifications	Description	Actions de reconfiguration
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
prestosql-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoSQL)

Classifications	Description	Actions de reconfiguration
prestosql-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	Modifiez les valeurs dans le fichier node.properties de PrestoSQL.	Not available.
prestosql-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de PrestoSQL.	Not available.
prestosql-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de PrestoSQL.	Not available.
prestosql-connector-hive	Modifiez les valeurs dans le fichier hive.properties de PrestoSQL.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de PrestoSQL.	Not available.
prestosql-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de PrestoSQL.	Not available.

Classifications	Description	Actions de reconfiguration
prestosql-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de PrestoSQL.	Not available.
prestosql-connector-memory	Modifiez les valeurs dans le fichier memory.properties de PrestoSQL.	Not available.
prestosql-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de PrestoSQL.	Not available.
prestosql-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de PrestoSQL.	Not available.
prestosql-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de PrestoSQL.	Not available.
prestosql-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de PrestoSQL.	Not available.
prestosql-connector-redis	Modifiez les valeurs dans le fichier redis.properties de PrestoSQL.	Not available.
prestosql-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de PrestoSQL.	Not available.
prestosql-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de PrestoSQL.	Not available.

Classifications	Description	Actions de reconfiguration
prestosql-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de PrestoSQL.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.

Classifications	Description	Actions de reconfiguration
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Amazon EMR version 6.2.0

Versions d'application 6.2.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, PrestoSQL, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.2.0	emr-6.1.1	emr-6.1.0	emr-6.0.1
AWS SDK pour Java	1,1880	1,1,828	1,1,828	1,1,711

	emr-6.2.0	emr-6.1.1	emr-6.1.0	emr-6.0.1
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,10	2,1,10	2,1,10	2,1,10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.0	1.11.0	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6-amzn-0	2.2.5	2.2.5	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.6.0-amzn-1	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1
Hue	4.8.0	4.7.1	4.7.1	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.0.0
Livy	0.7.0	0.7.0	0.7.0	0.6.0
MXNet	1.7.0	1.6.0	1.6.0	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.2.0	5.1.0

	emr-6.2.0	emr-6.1.1	emr-6.1.0	emr-6.0.1
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	-
Presto	0,238,3	0.232	0.232	0.230
Spark	3.0.1	3.0.0	3.0.0	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	-
TensorFlow	2.3.1	2.1.0	2.1.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	343	338	338	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 6.2.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.2.0. Les modifications ont été apportées à la version 6.1.0.

Date de parution initiale : 9 décembre 2020

Dernière mise à jour : 4 octobre 2021

Applications prises en charge

- AWS SDK pour Java version 1.11.828
- emr-record-server version 1.7.0
- Flink version 1.11.2
- Ganglia version 3.7.2
- Hadoop version 3.2.1-amzn-1
- HBase version 2.2.6-amzn-0

- HBase-operator-tools 1.0.0
- HCatalog version 3.1.2-amzn-0
- Hive version 3.1.2-amzn-3
- Hudi version 0.6.0-amzn-1
- Hue version 4.8.0
- JupyterHub version 1.1.0
- Livy version 0.7.0
- MXNet version 1.7.0
- Oozie version 5.2.0
- Phoenix version 5.0.0
- Pig version 0.17.0
- Presto version 0.238.3-amzn-1
- PrestoSQL version 343
- Spark version 3.0.1-amzn-0
- spark-rapids 0.2.0
- TensorFlow version 2.3.1
- Zeppelin version 0.9.0-preview1
- Zookeeper version 3.4.14
- Connecteurs et pilotes : Connecteur DynamoDB 4.16.0

Nouvelles fonctionnalités

- HBase: Suppression du changement de nom lors de la phase de validation et ajout d'un HFile suivi persistant. Consultez la section HFile Suivi permanent dans le guide de mise à jour d'Amazon EMR.
- HBase: Rétroporté Créez une configuration qui oblige à mettre en cache les blocs lors du compactage.
- PrestoDB : améliorations apportées à l'élimination dynamique des partitions. Join Reorder basée sur des règles fonctionne sur des données non partitionnées.
- Politiques gérées délimitées : pour s'aligner sur les AWS meilleures pratiques, Amazon EMR a introduit des politiques gérées par défaut définies dans la version 2 EMR en remplacement des politiques qui seront déconseillées. Consultez Politiques gérées par Amazon EMR.

- État de prise en charge du service de métadonnées d'instance (IMDS) V2 : pour Amazon EMR 6.2 ou version ultérieure, les composants Amazon EMR sont IMDSv2 utilisés pour tous les appels IMDS. Pour les appels IMDS dans le code de votre application, vous pouvez utiliser les deux IMDSv1 ou configurer l'IMDS pour qu'il ne soit utilisé que IMDSv2 pour renforcer la sécurité. IMDSv2 Si vous la désactivez IMDSv1 dans les versions antérieures d'Amazon EMR 6.x, cela entraîne un échec du démarrage du cluster.

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.
- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon

EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.

- Spark : amélioration des performances dans l'environnement d'exécution de Spark.

Problèmes connus

- Les autorisations définies sur Amazon EMR 6.2 sont incorrectes. Le répertoire/etc/cron.d/libinstance-controller-java file in EMR 6.2.0. Permissions on the file are 645 (-rw-r--r-x), when they should be 644 (-rw-r--r--). As a result, Amazon EMR version 6.2 does not log instance-state logs, and the /emr/instance-logs est vide. Ce problème est corrigé dans Amazon EMR 6.3.0 et les versions ultérieures.

  Pour contourner ce problème, exécutez le script suivant en tant qu'action d'amorçage lors du lancement du cluster.

  ```
  #!/bin/bash
  sudo chmod 644 /etc/cron.d/libinstance-controller-java
  ```

- Pour les clusters de sous-réseaux privés Amazon EMR 6.2.0 et 6.3.0, vous ne pouvez pas accéder à l'interface utilisateur Web de Ganglia. Vous recevrez un message d'erreur « accès refusé (403) ». D'autres sites Web UIs, tels que Spark, Hue JupyterHub, Zeppelin, Livy et Tez, fonctionnent normalement. L'accès à l'interface utilisateur Web de Ganglia sur les clusters de sous-réseaux publics fonctionne également normalement. Pour résoudre ce problème, redémarrez le service httpd sur le nœud primaire avec `sudo systemctl restart httpd`. Ce problème est résolu dans Amazon EMR 6.4.0.
- Amazon EMR 6.2.0 présente un problème selon lequel httpd échoue continuellement, ce qui rend Ganglia indisponible. Le message d'erreur « Impossible de se connecter au serveur » s'affiche. Pour réparer un cluster déjà en cours d'exécution présentant ce problème, connectez-vous en SSH au nœud primaire du cluster et ajoutez la ligne `Listen 80` au fichier `httpd.conf` situé dans `/etc/httpd/conf/httpd.conf`. Ce problème est résolu dans Amazon EMR 6.3.0.
- HTTTD échoue sur les clusters EMR 6.2.0 lorsque vous utilisez une configuration de sécurité. Cela rend l'interface utilisateur de l'application web Ganglia indisponible. Pour accéder à l'interface utilisateur de l'application web Ganglia, ajoutez `Listen 80` au fichier `/etc/httpd/conf/httpd.conf` sur le nœud primaire de votre cluster. Pour plus d'informations sur la connexion à votre cluster, consultez Connexion au nœud primaire à l'aide de SSH.

Blocs-notes EMR ne parvient pas non plus à établir une connexion avec les clusters EMR 6.2.0 lorsque vous utilisez une configuration de sécurité. Le bloc-notes ne parviendra pas à répertorier les noyaux et à soumettre les tâches Spark. Nous vous recommandons d'utiliser Blocs-notes EMR avec une autre version d'Amazon EMR à la place.

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Redémarrer InstanceController

     ```
     $ sudo systemctl daemon-reload

     $ sudo systemctl restart instance-controller
     ```

Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Amazon EMR 6.1.0 et 6.2.0 présentent un problème de performance qui peut affecter de manière critique toutes les opérations insert, upsert et delete de Hudi. Si vous envisagez d'utiliser Hudi avec Amazon EMR 6.1.0 ou 6.2.0, AWS contactez le support pour obtenir un RPM Hudi corrigé.

- **Important**

  Les clusters EMR qui exécutent Amazon Linux ou Amazon Linux 2 Amazon Machine Images (AMIs) utilisent le comportement par défaut d'Amazon Linux et ne téléchargent ni n'installent automatiquement les mises à jour importantes et critiques du noyau nécessitant un redémarrage. Ce comportement est identique à celui des autres EC2 instances Amazon qui exécutent l'AMI Amazon Linux par défaut. Si de nouvelles mises à jour logicielles Amazon Linux nécessitant un redémarrage (telles que les mises à jour du noyau, de

NVIDIA et de CUDA) sont disponibles après la publication d'une version d'Amazon EMR, les instances de cluster EMR qui exécutent l'AMI par défaut ne téléchargent pas et n'installent pas automatiquement ces mises à jour. Pour obtenir les mises à jour du noyau, vous pouvez personnaliser votre AMI Amazon EMR afin d'utiliser la dernière AMI Amazon Linux.

- Les artefacts Maven d'Amazon EMR 6.2.0 ne sont pas publiés. Ils seront publiés avec une future version d'Amazon EMR.
- Le HFile suivi permanent à l'aide de la table système HBase Storefile ne prend pas en charge la fonctionnalité de réplication des HBase régions. Pour plus d'informations sur HBase la réplication régionale, consultez la section Nombre élevé de lectures disponibles cohérentes avec la chronologie.
- Différences de version entre Amazon EMR 6.x et EMR 5.x pour la compartimentation Hive

  EMR 5.x utilise OOS Apache Hive 2, tandis que EMR 6.x utilise OOS Apache Hive 3. La version open source Hive2 utilise la version 1 de Bucketing, tandis que la version open source Hive3 utilise la version 2. Cette différence de version de compartimentation entre Hive 2 (EMR 5.x) et Hive 3 (EMR 6.x) signifie que le hachage de compartimentation de Hive fonctionne différemment. Consultez l'exemple ci-dessous.

  Le tableau suivant est un exemple créé dans EMR 6.x et EMR 5.x, respectivement.

```
-- Using following LOCATION in EMR 6.x
CREATE TABLE test_bucketing (id INT, desc STRING)
PARTITIONED BY (day STRING)
CLUSTERED BY(id) INTO 128 BUCKETS
LOCATION 's3://your-own-s3-bucket/emr-6-bucketing/';

-- Using following LOCATION in EMR 5.x
LOCATION 's3://your-own-s3-bucket/emr-5-bucketing/';
```

  Insertion des mêmes données dans EMR 6.x et EMR 5.x.

```
INSERT INTO test_bucketing PARTITION (day='01') VALUES(66, 'some_data');
INSERT INTO test_bucketing PARTITION (day='01') VALUES(200, 'some_data');
```

  La vérification de l'emplacement S3 montre que le nom du fichier de compartimentation est différent, car la fonction de hachage est différente entre EMR 6.x (Hive 3) et EMR 5.x (Hive 2).

```
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-6-bucketing/day=01/
2020-10-21 20:35:16         13 000025_0
2020-10-21 20:35:22         14 000121_0
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-5-bucketing/day=01/
2020-10-21 20:32:07         13 000066_0
2020-10-21 20:32:51         14 000072_0
```

Vous pouvez également constater la différence de version en exécutant la commande suivante dans la CLI Hive dans EMR 6.x. Notez qu'il renvoie la version 2 de compartimentation.

```
hive> DESCRIBE FORMATTED test_bucketing;
...
Table Parameters:
    bucketing_version       2
...
```

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 6.2.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	3.1.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.0.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.16.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.0.0	Connecteur S3Select EMR
emrfs	2,44,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.11.2	Applications et scripts client de la ligne de commande Apache Flink.

Composant	Version	Description
flink-jobmanager-config	1.11.2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.2.1-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.2.1-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-2	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.

Composant	Version	Description
hadoop-httpfs-server	3.2.1-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.2.1-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.2.1-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.2.6-amzn-0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.2.6-amzn-0	Service pour desservir une ou plusieurs HBase régions.
hbase-client	2.2.6-amzn-0	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	2.2.6-amzn-0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.2.6-amzn-0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	3.1.2-amzn-3	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.2-amzn-3	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.2-amzn-3	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.2-amzn-3	Client de ligne de commande Hive.
hive-hbase	3.1.2-amzn-3	Client Hive-hbase.
hive-metastore-server	3.1.2-amzn-3	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.2-amzn-3	Service pour l'acceptation de requêtes Hive en tant que requêtes web.

Composant	Version	Description
hudi	0.6.0-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.6.0-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-prestosql	0.6.0-amzn-1	Bibliothèque de solution groupée pour exécuter PrestoSQL avec Hudi.
hudi-spark	0.6.0-amzn-1	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.8.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.1.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.7.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5.5.64+	Serveur de base de données MariaDB.

Composant	Version	Description
nvidia-cuda	10,1243	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5,0,0- -2,0 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	5,0,0- -2,0 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.238.3-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.238.3-amzn-1	Service pour exécuter les éléments d'une requête.
presto-client	0.238.3-amzn-1	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.

Composant	Version	Description
prestosql-coordinator	343	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants prestosql-workers.
prestosql-worker	343	Service pour exécuter les éléments d'une requête.
prestosql-client	343	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.3	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.0.1-amzn-0	Clients de ligne de commande Spark.
spark-history-server	3.0.1-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.0.1-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.0.1-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.

Composant	Version	Description
spark-rapids	0.2.0	Plugin Nvidia Spark RAPIDS qui accélère Apache Spark avec GPUs.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.3.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'application tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0.9.0-preview1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 6.2.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-6.2.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Not available.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Not available.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.	Not available.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	This classification should not be reconfigured.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.	Restarts Hadoop HDFS ZKFC.

Classifications	Description	Actions de reconfiguration
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat server.
hive	Paramètres définis par Amazon EMR pour Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.

Classifications	Description	Actions de reconfiguration
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.

Classifications	Description	Actions de reconfiguration
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
prestosql-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server (for PrestoSQL)

Classifications	Description	Actions de reconfiguration
prestosql-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	Modifiez les valeurs dans le fichier node.properties de PrestoSQL.	Not available.
prestosql-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de PrestoSQL.	Not available.
prestosql-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de PrestoSQL.	Not available.
prestosql-connector-hive	Modifiez les valeurs dans le fichier hive.properties de PrestoSQL.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de PrestoSQL.	Not available.
prestosql-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de PrestoSQL.	Not available.

Classifications	Description	Actions de reconfiguration
prestosql-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de PrestoSQL.	Not available.
prestosql-connector-memory	Modifiez les valeurs dans le fichier memory.properties de PrestoSQL.	Not available.
prestosql-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de PrestoSQL.	Not available.
prestosql-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de PrestoSQL.	Not available.
prestosql-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de PrestoSQL.	Not available.
prestosql-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de PrestoSQL.	Not available.
prestosql-connector-redis	Modifiez les valeurs dans le fichier redis.properties de PrestoSQL.	Not available.
prestosql-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de PrestoSQL.	Not available.
prestosql-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de PrestoSQL.	Not available.

Classifications	Description	Actions de reconfiguration
prestosql-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de PrestoSQL.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.

Classifications	Description	Actions de reconfiguration
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restart Oozie.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Amazon EMR version 6.1.1

Versions d’application 6.1.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, PrestoSQL, Spark, Sqoop, TensorFlow, TezZeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK pour Java	1,1,828	1,1,828	1,1,711	1,1,711

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,10	2,1,10	2,1,10	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 6.1.1

Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.

Modifications, améliorations et problèmes résolus

- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.

- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Le protocole HTTPS est désormais activé par défaut pour les référentiels Amazon Linux. Si vous utilisez une politique VPCE Amazon S3 pour restreindre l'accès à des compartiments spécifiques, vous devez ajouter le nouvel ARN du compartiment Amazon Linux `arn:aws:s3:::amazonlinux-2-repos-$region/*` à votre politique (remplacez `$region` par la région où se trouve le point de terminaison). Pour plus d'informations, consultez cette rubrique dans les forums de AWS discussion. Annonce : Amazon Linux 2 permet désormais d'utiliser le protocole HTTPS lors de la connexion aux référentiels de packages.

Versions des composants 6.1.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*`-amzn-`*EmrVersion*. *EmrVersion*

commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.3.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.14.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	3.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.14.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	2.0.0	Connecteur S3Select EMR
emrfs	2,42,0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.11.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.2.1-amzn-1.1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-1.1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-1.1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.2.1-amzn-1.1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-1.1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.2.1-amzn-1.1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-1.1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	3.2.1-amzn-1.1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.2.1-amzn-1.1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.2.1-amzn-1.1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-1.1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.2.5	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.2.5	Service pour desservir une ou plusieurs HBase régions.
hbase-client	2.2.5	HBase client en ligne de commande.
hbase-rest-server	2.2.5	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.2.5	Service fournissant un point de terminaison Thrift pour. HBase

Composant	Version	Description
hcatalog-client	3.1.2-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.2-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.2-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.2-amzn-2	Client de ligne de commande Hive.
hive-hbase	3.1.2-amzn-2	Client Hive-hbase.
hive-metastore-server	3.1.2-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.2-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.5.2-incubating-amzn-2	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.5.2-incubating-amzn-2	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.

Composant	Version	Description
hudi-prestosql	0.5.2-incubating-amzn-2	Bibliothèque de solution groupée pour exécuter PrestoSQL avec Hudi.
hudi-spark	0.5.2-incubating-amzn-2	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.7.1	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.1.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.6.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5.5.64+	Serveur de base de données MariaDB.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.

Composant	Version	Description
opencv	4.3.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5,0,0- -2,0 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	5,0,0- -2,0 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.232	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.232	Service pour exécuter les éléments d'une requête.
presto-client	0.232	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
prestosql-coordinator	338	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants prestosql-workers.
prestosql-worker	338	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
prestosql-client	338	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.3	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.0.0-amzn-0.1	Clients de ligne de commande Spark.
spark-history-server	3.0.0-amzn-0.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.0.0-amzn-0.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.0.0-amzn-0.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.1.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.

Composant	Version	Description
tez-on-yarn	0.9.2	Les bibliothèques et l'application tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0.9.0-preview1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 6.1.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-6.1.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.

Classifications	Description
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.

Classifications	Description
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive	Paramètres définis par Amazon EMR pour Apache Hive.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-l og4j2.properties de Hive.

Classifications	Description
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.

Classifications	Description
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.

Classifications	Description
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.

Classifications	Description
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
prestosql-log	Modifiez les valeurs dans le fichier log.properties de Presto.
prestosql-config	Modifiez les valeurs dans le fichier config.properties de Presto.
prestosql-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
prestosql-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
prestosql-node	Modifiez les valeurs dans le fichier node.properties de PrestoSQL.
prestosql-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de PrestoSQL.
prestosql-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de PrestoSQL.
prestosql-connector-hive	Modifiez les valeurs dans le fichier hive.properties de PrestoSQL.
prestosql-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de PrestoSQL.
prestosql-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de PrestoSQL.
prestosql-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de PrestoSQL.
prestosql-connector-memory	Modifiez les valeurs dans le fichier memory.properties de PrestoSQL.

Classifications	Description
prestosql-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de PrestoSQL.
prestosql-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de PrestoSQL.
prestosql-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de PrestoSQL.
prestosql-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de PrestoSQL.
prestosql-connector-redis	Modifiez les valeurs dans le fichier redis.properties de PrestoSQL.
prestosql-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de PrestoSQL.
prestosql-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de PrestoSQL.
prestosql-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de PrestoSQL.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.

Classifications	Description
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Classifications	Description
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 6.1.0

Versions d’application 6.1.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, Presto, PrestoSQL, Spark, Sqoop, TensorFlow, TezZeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK pour Java	1,1,828	1,1,828	1,1,711	1,1,711
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,10	2,1,10	2,1,10	2,1,12

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 6.1.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.1.0. Les modifications ont été apportées à la version 6.0.0.

Date de parution initiale : 4 septembre 2020

Dernière mise à jour : 15 octobre 2020

Applications prises en charge

- AWS SDK pour Java version 1.11.828
- Flink version 1.11.0
- Ganglia version 3.7.2
- Hadoop version 3.2.1-amzn-1
- HBase version 2.2.5
- HBase-operator-tools 1.0.0
- HCatalog version 3.1.2-amzn-0
- Hive version 3.1.2-amzn-1

- Hudi version 0.5.2-incubation
- Hue version 4.7.1
- JupyterHub version 1.1.0
- Livy version 0.7.0
- MXNet version 1.6.0
- Oozie version 5.2.0
- Phoenix version 5.0.0
- Presto version 0.232
- PrestoSQL version 338
- Spark version 3.0.0-amzn-0
- TensorFlow version 2.1.0
- Zeppelin version 0.9.0-preview1
- Zookeeper version 3.4.14
- Connecteurs et pilotes : Connecteur DynamoDB 4.14.0

Nouvelles fonctionnalités

- Les types d'instance ARM sont pris en charge à partir des versions 5.30.0 et 6.1.0 d'Amazon EMR.
- Les types d'instances à usage général M6g sont pris en charge à partir des versions 6.1.0 et 5.30.0 d'Amazon EMR. Pour plus d'informations, consultez Types d'instance pris en charge dans le Guide de gestion Amazon EMR.
- La fonctionnalité de groupe de EC2 placement est prise en charge à partir de la version 5.23.0 d'Amazon EMR en tant qu'option pour plusieurs clusters de nœuds principaux. Actuellement, seuls les types de nœuds primaires sont pris en charge par la fonctionnalité de groupe de placement. La stratégie `SPREAD` est appliquée à ces nœuds primaires. La stratégie `SPREAD` place un petit groupe d'instances sur un matériel sous-jacent distinct afin de se prémunir contre la perte de plusieurs nœuds primaires en cas de panne matérielle. Pour plus d'informations, consultez la section Intégration de l'EMR au groupe de EC2 placement dans le guide de gestion Amazon EMR.
- Mise à l'échelle gérée – Avec la version 6.1.0 d'Amazon EMR, vous pouvez activer la mise à l'échelle gérée par Amazon EMR pour augmenter ou diminuer automatiquement le nombre d'instances ou d'unités dans votre cluster en fonction de la charge de travail. Amazon EMR évalue en permanence les métriques de cluster pour prendre des décisions de dimensionnement qui optimisent vos clusters en termes de coût et de vitesse. Managed Scaling est également disponible

sur Amazon EMR version 5.30.0 et suivantes, à l'exception de 6.0.0. Pour plus d'informations, consultez Mise à l'échelle des ressources de cluster dans le Guide de gestion Amazon EMR.

- La version 338 de PrestoSQL est prise en charge avec EMR 6.1.0. Pour plus d'informations, consultez Presto.
  - PrestoSQL est uniquement pris en charge sur EMR 6.1.0 et versions ultérieures, pas sur EMR 6.0.0 ou EMR 5.x.
  - Le nom de l'application, `Presto` continue d'être utilisé pour installer PrestoDB sur des clusters. Pour installer PrestoSQL sur des clusters, utilisez le nom de l'application `PrestoSQL`.
  - Vous pouvez installer PrestoDB ou PrestoSQL, mais vous ne pouvez pas installer les deux sur un seul cluster. Si PrestoDB et PrestoSQL sont spécifiés lors de la tentative de création d'un cluster, une erreur de validation se produit et la demande de création de cluster échoue.
  - PrestoSQL est pris en charge sur les clusters mono-maître et muti-maître. Sur les clusters multi-maîtres, un métastore Hive externe est requis pour exécuter PrestoSQL ou Prestodb. Consultez Applications prises en charge dans un cluster EMR avec plusieurs nœuds primaires.
- Prise en charge de l'authentification automatique ECR sur Apache Hadoop et Apache Spark avec Docker : les utilisateurs de Spark peuvent utiliser des images Docker provenant de Docker Hub et d'Amazon Elastic Container Registry (Amazon ECR) pour définir les dépendances de l'environnement et des bibliothèques.

  Configurez Docker et exécutez des applications Spark avec Docker à l'aide d'Amazon EMR 6.x.
- EMR prend en charge les transactions ACID d'Apache Hive : Amazon EMR 6.1.0 prend en charge les transactions ACID de Hive afin qu'elles soient conformes aux propriétés ACID d'une base de données. Grâce à cette fonctionnalité, vous pouvez exécuter les opérations `INSERT, UPDATE, DELETE,` et `MERGE` dans des tables gérées par Hive avec des données stockées dans Amazon Simple Storage Service (Amazon S3). Il s'agit d'une fonctionnalité essentielle pour les cas d'utilisation tels que l'ingestion en continu, le retraitement des données, les mises à jour en masse à l'aide de MERGE et les dimensions qui changent lentement. Pour plus d'informations, notamment des exemples de configuration et des cas supports, consultez Amazon EMR prend en charge les transactions ACID d'Apache Hive.

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.

- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Apache Flink n'est pas pris en charge sur EMR 6.0.0, mais il l'est sur EMR 6.1.0 avec Flink 1.11.0. Il s'agit de la première version de Flink officiellement compatible avec Hadoop 3. Consultez Annonce de sortie d'Apache Flink 1.11.0.
- Ganglia a été supprimé des packages EMR 6.1.0 par défaut.

Problèmes connus

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon

Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

Définir explicitement un ulimit à partir de la ligne de commande

1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Redémarrer InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
```

```
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- > **Important**
  >
  > Amazon EMR 6.1.0 et 6.2.0 présentent un problème de performance qui peut affecter de manière critique toutes les opérations insert, upsert et delete de Hudi. Si vous envisagez d'utiliser Hudi avec Amazon EMR 6.1.0 ou 6.2.0, AWS contactez le support pour obtenir un RPM Hudi corrigé.

- Si vous définissez une configuration de collecte des déchets personnalisée avec `spark.driver.extraJavaOptions` et`spark.executor.extraJavaOptions`, cela entraînera un échec du driver/executor lancement d'EMR 6.1 en raison d'une configuration de collecte des déchets conflictuelle. Avec la version 6.1.0 d'EMR, vous devez spécifier la configuration personnalisée du récupérateur de mémoire Spark pour les pilotes et les exécuteurs avec les propriétés `spark.driver.defaultJavaOptions` et `spark.executor.defaultJavaOptions` à la place. Pour en savoir plus, consultez Environnement d'exécution Apache Spark et Configuration du récupérateur de mémoire Spark sur Amazon EMR 6.1.0.
- L'utilisation de Pig avec Oozie (et au sein de Hue, puisque Hue utilise les actions Oozie pour exécuter les scripts Pig), génère une erreur indiquant qu'une bibliothèque native-lzo ne peut pas être chargée. Ce message d'erreur est informatif et n'empêche pas Pig de fonctionner.
- Prise en charge de la simultanéité dans Hudi : Actuellement, Hudi ne prend pas en charge les écritures simultanées dans une seule table Hudi. De plus, Hudi annule toutes les modifications effectuées par les enregistreurs en cours avant de permettre à un nouvel enregistreur de démarrer. Les écritures simultanées peuvent interférer avec ce mécanisme et introduire des conditions de

concurrence, ce qui peut entraîner une corruption des données. Vous devez vous assurer que, dans le cadre de votre flux de traitement des données, il n'y a qu'un seul enregistreur Hudi opérant sur une table Hudi à tout moment. Hudi prend en charge plusieurs lecteurs simultanés opérant sur la même table Hudi.

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

- Il y a un problème dans Amazon EMR 6.1.0 qui affecte les clusters exécutant Presto. Après une période prolongée (plusieurs jours), le cluster peut générer des erreurs telles que "su : failed to execute /bin/bash : Resource temporarily unavailable" ou "shell request failed on channel 0". Ce problème est dû à un processus interne Amazon EMR (InstanceController) qui génère un trop grand nombre de processus légers (LWP), ce qui finit par amener l'utilisateur Hadoop à dépasser

sa limite nproc. Cela empêche l'utilisateur d'ouvrir des processus supplémentaires. La solution à ce problème est de mettre à niveau vers EMR 6.2.0.

Versions des composants 6.1.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.3.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.14.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	3.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.14.0	Application de copie distribuée optimisée pour Amazon S3.

Composant	Version	Description
emr-s3-select	2.0.0	Connecteur S3Select EMR
emrfs	2,42,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.11.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.2.1-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-1	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	3.2.1-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.2.1-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.2.1-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.2.1-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	2.2.5	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.2.5	Service pour desservir une ou plusieurs HBase régions.
hbase-client	2.2.5	HBase client en ligne de commande.
hbase-rest-server	2.2.5	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.2.5	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	3.1.2-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.2-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.2-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.2-amzn-2	Client de ligne de commande Hive.
hive-hbase	3.1.2-amzn-2	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	3.1.2-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.2-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.5.2-incubating-amzn-2	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.5.2-incubating-amzn-2	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hudi-prestosql	0.5.2-incubating-amzn-2	Bibliothèque de solution groupée pour exécuter PrestoSQL avec Hudi.
hudi-spark	0.5.2-incubating-amzn-2	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hue-server	4.7.1	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.1.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark

Composant	Version	Description
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.6.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5.5.64+	Serveur de base de données MariaDB.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.3.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5,0,0- -2,0 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	5,0,0- -2,0 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.232	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.232	Service pour exécuter les éléments d'une requête.
presto-client	0.232	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
prestosql-coordinator	338	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants prestosql-workers.
prestosql-worker	338	Service pour exécuter les éléments d'une requête.
prestosql-client	338	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.3	Projet R pour les calculs statistiques
ranger-kms-server	2.0.0	Système de gestion des clés Apache Ranger
spark-client	3.0.0-amzn-0	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	3.0.0-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	3.0.0-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	3.0.0-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.1.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0.9.0-preview1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 6.1.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-6.1.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.

Classifications	Description
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.

Classifications	Description
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive	Paramètres définis par Amazon EMR pour Apache Hive.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.

Classifications	Description
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
prestosql-log	Modifiez les valeurs dans le fichier log.properties de Presto.
prestosql-config	Modifiez les valeurs dans le fichier config.properties de Presto.
prestosql-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
prestosql-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
prestosql-node	Modifiez les valeurs dans le fichier node.properties de PrestoSQL.

Classifications	Description
prestosql-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de PrestoSQL.
prestosql-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de PrestoSQL.
prestosql-connector-hive	Modifiez les valeurs dans le fichier hive.properties de PrestoSQL.
prestosql-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de PrestoSQL.
prestosql-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de PrestoSQL.
prestosql-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de PrestoSQL.
prestosql-connector-memory	Modifiez les valeurs dans le fichier memory.properties de PrestoSQL.
prestosql-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de PrestoSQL.
prestosql-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de PrestoSQL.
prestosql-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de PrestoSQL.
prestosql-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de PrestoSQL.
prestosql-connector-redis	Modifiez les valeurs dans le fichier redis.properties de PrestoSQL.
prestosql-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de PrestoSQL.

Classifications	Description
prestosql-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de PrestoSQL.
prestosql-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de PrestoSQL.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.

Classifications	Description
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 6.0.1

Versions d’application 6.0.1

Cette version inclut les applications suivantes : GangliaHBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterHub, Livy, MXNet, Oozie, PhoenixPresto, Spark, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK pour Java	1,1,828	1,1,828	1,1,711	1,1,711
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,10	2,1,10	2,1,10	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 6.0.1

Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.

Modifications, améliorations et problèmes résolus

- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Le protocole HTTPS est désormais activé par défaut pour les référentiels Amazon Linux. Si vous utilisez une politique VPCE Amazon S3 pour restreindre l'accès à des compartiments spécifiques, vous devez ajouter le nouvel ARN du compartiment Amazon Linux `arn:aws:s3:::amazonlinux-2-repos-$region/*` à votre politique (remplacez `$region` par la région où se trouve le point de terminaison). Pour plus d'informations, consultez cette rubrique dans les forums de AWS discussion. Annonce : Amazon Linux 2 permet désormais d'utiliser le protocole HTTPS lors de la connexion aux référentiels de packages.

Versions des composants 6.0.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.6	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.14.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	3.0.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.14.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.5.0	Connecteur S3Select EMR

Composant	Version	Description
emrfs	2.39.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.2.1-amzn-0.1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-0.1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-0.1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	3.2.1-amzn-0.1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-0.1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.

Composant	Version	Description
hadoop-httpfs-server	3.2.1-amzn-0.1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-0.1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-0.1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.2.1-amzn-0.1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.2.1-amzn-0.1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-0.1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	2.2.3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.2.3	Service pour desservir une ou plusieurs HBase régions.
hbase-client	2.2.3	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	2.2.3	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.2.3	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	3.1.2-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.2-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.2-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.2-amzn-0	Client de ligne de commande Hive.
hive-hbase	3.1.2-amzn-0	Client Hive-hbase.
hive-metastore-server	3.1.2-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.2-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.

Composant	Version	Description
hudi	0.5.0-incubating-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.5.0-incubating-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.4.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	1.0.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.6.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.5.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5.5.64+	Serveur de base de données MariaDB.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.1.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5,0,0- -2,0 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	5,0,0- -2,0 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.230	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.230	Service pour exécuter les éléments d'une requête.
presto-client	0.230	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
r	3.4.3	Projet R pour les calculs statistiques
spark-client	2.4.4	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.4.4	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.4	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.4	Bibliothèques Apache Spark requises par les esclaves YARN.
tensorflow	1.14.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0.9.0-SNAPSHOT	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 6.0.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-6.0.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.

Classifications	Description
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.

Classifications	Description
hive	Paramètres définis par Amazon EMR pour Apache Hive.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.

Classifications	Description
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-note s Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.

Classifications	Description
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password- authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-en v.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.prop erties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra .properties de Presto.

Classifications	Description
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.

Classifications	Description
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site .xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.

Classifications	Description
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 6.0.0

Versions d'application 6.0.0

Cette version inclut les applications suivantes : GangliaHBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterHub, Livy, MXNet, Oozie, PhoenixPresto, Spark, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR

- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK pour Java	1,1,828	1,1,828	1,1,711	1,1,711
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,10	2,1,10	2,1,10	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 6.0.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.0.0.

Date de parution initiale : 10 mars 2020

Applications prises en charge

- AWS SDK pour Java version 1.11.711
- Ganglia version 3.7.2

- Hadoop version 3.2.1
- HBase version 2.2.3
- HCatalog version 3.1.2
- Hive version 3.1.2
- Hudi version 0.5.0-incubation
- Hue version 4.4.0
- JupyterHub version 1.0.0
- Livy version 0.6.0
- MXNet version 1.5.1
- Oozie version 5.1.0
- Phoenix version 5.0.0
- Presto version 0.230
- Spark version 2.4.4
- TensorFlow version 1.14.0
- Zeppelin version 0.9.0-SNAPSHOT
- Zookeeper version 3.4.14
- Connecteurs et pilotes : Connecteur DynamoDB 4.14.0

> **Note**
>
> Flink, Sqoop, Pig et Mahout ne sont pas disponibles dans Amazon EMR version 6.0.0.

Nouvelles fonctionnalités

- Prise en charge de l'exécution de Docker pour YARN - Les applications YARN, telles que les tâches Spark, peuvent désormais s'exécuter dans le contexte d'un conteneur Docker. Vous pouvez ainsi définir facilement des dépendances dans une image Docker sans avoir besoin d'installer des bibliothèques personnalisées sur votre cluster Amazon EMR. Pour plus d'informations, consultez Configuration de l'intégration Docker et Exécution des applications Spark avec Docker à l'aide d'Amazon EMR version 6.0.0.

- Prise en charge de LLAP pour Hive - Hive prend désormais en charge le mode d'exécution LLAP pour améliorer les performances des requêtes. Pour de plus amples informations, veuillez consulter Utilisation de LLAP pour Hive.

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.
- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Amazon Linux
  - Amazon Linux 2 est le système d'exploitation de la série EMR version 6.x.

  - `systemd` est utilisé pour la gestion des services au lieu d'`upstart`, utilisé dans Amazon Linux 1.
- Kit de développement Java (JDK)
  - Corretto JDK 8 est le JDK par défaut pour la série EMR version 6.x.
- Scala
  - Scala 2.12 est utilisé avec Apache Spark et Apache Livy.
- Python 3
  - Python 3 est maintenant la version par défaut de Python dans EMR.
- Étiquettes de nœud YARN
  - À partir de la série Amazon EMR version 6.x, la fonction des étiquettes de nœud YARN est désactivée par défaut. Les processus principaux des applications peuvent s'exécuter à la fois sur les nœuds de noyau et sur les nœuds de tâche par défaut. Vous pouvez activer la fonction des étiquettes de nœud YARN en configurant les propriétés suivantes : `yarn.node-labels.enabled` et `yarn.node-labels.am.default-node-label-expression`. Pour plus d'informations, consultez Comprendre les nœuds primaires, les nœuds principaux et les nœuds de tâches.

Problèmes connus

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

Définir explicitement un ulimit à partir de la ligne de commande

1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Redémarrer InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
```

```
sudo systemctl daemon-reload
```

- Le shell interactif Spark PySpark, y compris SparkR et spark-shell, ne prend pas en charge l'utilisation de Docker avec des bibliothèques supplémentaires.
- Pour utiliser Python 3 avec Amazon EMR version 6.0.0, vous devez ajouter `PATH` à `yarn.nodemanager.env-whitelist`.
- La fonctionnalité Live Long and Process (LLAP) n'est pas prise en charge lorsque vous utilisez le catalogue de données AWS Glue comme métastore pour Hive.
- Lorsque vous utilisez Amazon EMR 6.0.0 avec l'intégration Spark et Docker, vous devez configurer les instances de votre cluster avec le même type d'instance et la même quantité de volumes EBS pour éviter les échecs lors de la soumission d'une tâche Spark avec l'exécution Docker.
- Dans Amazon EMR 6.0.0, le mode de stockage Amazon HBase S3 est impacté par le problème HBASE-24286. HBase master ne peut pas s'initialiser lorsque le cluster est créé à l'aide de données S3 existantes.
- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 6.0.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.6	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.14.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	3.0.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.14.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.5.0	Connecteur S3Select EMR
emrfs	2.39.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	3.2.1-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	3.2.1-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	3.2.1-amzn-0	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	3.2.1-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	3.2.1-amzn-0	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	3.2.1-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	3.2.1-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	3.2.1-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	3.2.1-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	3.2.1-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	2.2.3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	2.2.3	Service pour desservir une ou plusieurs HBase régions.
hbase-client	2.2.3	HBase client en ligne de commande.
hbase-rest-server	2.2.3	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	2.2.3	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	3.1.2-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	3.1.2-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	3.1.2-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	3.1.2-amzn-0	Client de ligne de commande Hive.
hive-hbase	3.1.2-amzn-0	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	3.1.2-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	3.1.2-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.5.0-incubating-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.5.0-incubating-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.4.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.0.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.6.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mxnet	1.5.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.

Composant	Version	Description
mariadb-server	5.5.64+	Serveur de base de données MariaDB.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	5,0,0- -2,0 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	5,0,0- -2,0 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.230	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.230	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
presto-client	0.230	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
r	3.4.3	Projet R pour les calculs statistiques
spark-client	2.4.4	Clients de ligne de commande Spark.
spark-history-server	2.4.4	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.4	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.4	Bibliothèques Apache Spark requises par les esclaves YARN.
tensorflow	1.14.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.41+	Serveur HTTP Apache.
zeppelin-server	0.9.0-SNAPSHOT	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 6.0.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-6.0.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-env	Modifiez les valeurs dans l'environnement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.

Classifications	Description
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive	Paramètres définis par Amazon EMR pour Apache Hive.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2

Classifications	Description
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password- authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-en v.sh de Presto.

Classifications	Description
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.

Classifications	Description
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.

Classifications	Description
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Versions Amazon EMR 5.x

Cette section contient les versions des applications, les notes de mise à jour, les versions des composants et les classifications de configuration disponibles dans chaque version 5.x d'Amazon EMR.

Lorsque vous lancez un cluster, vous pouvez choisir parmi plusieurs versions d'Amazon EMR. Vous pouvez ainsi tester et utiliser des versions d'applications qui répondent à vos exigences de compatibilité. Vous spécifiez le numéro de version avec l'étiquette de version. Les étiquettes de version sont sous la forme `emr-x.x.x`. Par exemple, `emr-7.9.0`.

Les nouvelles versions d'Amazon EMR sont mises à disposition dans différentes régions sur une période de plusieurs jours, en commençant par la première région à la date de sortie initiale. Il est possible que la dernière version ne soit pas disponible dans votre région pendant cette période.

Pour un tableau complet des versions des applications dans chaque version 5.x d'Amazon EMR, consultez Versions des applications dans les versions 5.x d'Amazon EMR.

Rubriques

- Versions des applications dans les versions 5.x d'Amazon EMR
- Amazon EMR version 5.36.2
- Amazon EMR version 5.36.1
- Amazon EMR version 5.36.0
- Amazon EMR version 5.35.0
- Amazon EMR version 5.34.0
- Amazon EMR version 5.33.1
- Amazon EMR version 5.33.0
- Amazon EMR version 5.32.1
- Amazon EMR version 5.32.0
- Amazon EMR version 5.31.1
- Amazon EMR version 5.31.0
- Amazon EMR version 5.30.2
- Amazon EMR version 5.30.1
- Amazon EMR version 5.30.0
- Amazon EMR version 5.29.0
- Amazon EMR version 5.28.1
- Amazon EMR version 5.28.0
- Amazon EMR version 5.27.1

- Amazon EMR version 5.27.0
- Amazon EMR version 5.26.0
- Amazon EMR version 5.25.0
- Amazon EMR version 5.24.1
- Amazon EMR version 5.24.0
- Amazon EMR version 5.23.1
- Amazon EMR version 5.23.0
- Amazon EMR version 5.22.0
- Amazon EMR version 5.21.2
- Amazon EMR version 5.21.1
- Amazon EMR version 5.21.0
- Amazon EMR version 5.20.1
- Amazon EMR version 5.20.0
- Amazon EMR version 5.19.1
- Amazon EMR version 5.19.0
- Amazon EMR version 5.18.1
- Amazon EMR version 5.18.0
- Amazon EMR version 5.17.2
- Amazon EMR version 5.17.1
- Amazon EMR version 5.17.0
- Amazon EMR version 5.16.1
- Amazon EMR version 5.16.0
- Amazon EMR version 5.15.1
- Amazon EMR version 5.15.0
- Amazon EMR version 5.14.2
- Amazon EMR version 5.14.1
- Amazon EMR version 5.14.0
- Amazon EMR version 5.13.1
- Amazon EMR version 5.13.0

- Amazon EMR version 5.12.3
- Amazon EMR version 5.12.2
- Amazon EMR version 5.12.1
- Amazon EMR version 5.12.0
- Amazon EMR version 5.11.4
- Amazon EMR version 5.11.3
- Amazon EMR version 5.11.2
- Amazon EMR version 5.11.1
- Amazon EMR version 5.11.0
- Amazon EMR version 5.10.1
- Amazon EMR version 5.10.0
- Amazon EMR version 5.9.1
- Amazon EMR version 5.9.0
- Amazon EMR version 5.8.3
- Amazon EMR version 5.8.2
- Amazon EMR version 5.8.1
- Amazon EMR version 5.8.0
- Amazon EMR version 5.7.1
- Amazon EMR version 5.7.0
- Amazon EMR version 5.6.1
- Amazon EMR version 5.6.0
- Amazon EMR version 5.5.4
- Amazon EMR version 5.5.3
- Amazon EMR version 5.5.2
- Amazon EMR version 5.5.1
- Amazon EMR version 5.5.0
- Amazon EMR version 5.4.1
- Amazon EMR version 5.4.0
- Amazon EMR version 5.3.2

- Amazon EMR version 5.3.1
- Amazon EMR version 5.3.0
- Amazon EMR version 5.2.3
- Amazon EMR version 5.2.2
- Amazon EMR version 5.2.1
- Amazon EMR version 5.2.0
- Amazon EMR version 5.1.1
- Amazon EMR version 5.1.0
- Amazon EMR version 5.0.3
- Amazon EMR version 5.0.2
- Amazon EMR version 5.0.1
- Amazon EMR version 5.0.0

Versions des applications dans les versions 5.x d'Amazon EMR

Pour obtenir un tableau complet répertoriant les versions d'applications disponibles dans chaque version d'Amazon EMR 5.x, ouvrez Versions des applications dans les versions 5.x d'Amazon EMR dans votre navigateur.

Amazon EMR version 5.36.2

versions de l'application 5.36.2

Cette version inclut les applications suivantes : DeltaFlinkGanglia, HBase, HCatalogHadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR

- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.36.2	emr-5.36.1	emr-5.36.0	emr-5.35.0
AWS SDK pour Java	1,12,206	1,12,206	1,12,206	1,12,159
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.2	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.9	2.3.9	2.3.9
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.9	2.3.9
Hudi	0.10.1-amzn-1	0.10.1-amzn-1	0.10.1-amzn-1	0.9.0-amzn-2
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.6.0	2.6.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1

	emr-5.36.2	emr-5.36.1	emr-5.36.0	emr-5.35.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,267	0,267	0,267	0,266
Spark	2.4.8	2.4.8	2.4.8	2.4.8
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.10.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

notes de mise à jour de la version 5.36.2

Les notes de mise à jour suivantes incluent des informations relatives à la version 5.36.2 d'Amazon EMR. Les modifications sont relatives à la version 5.36.1. Pour plus d'informations sur le calendrier de publication, consultez le journal des modifications.

Modifications, améliorations et problèmes résolus

- Cette version améliore la logique de réduction du cluster afin qu'Amazon EMR ne réduise pas les nœuds principaux en dessous du paramètre de facteur de réplication HDFS pour le cluster. Cette

amélioration répond aux exigences de redondance des données et réduit le risque de blocage d'une opération de dimensionnement.

- Cette version ajoute un nouveau mécanisme de nouvelle tentative au flux de travail de dimensionnement des clusters pour les clusters EMR qui exécutent Presto ou Trino. Cette amélioration réduit le risque que le redimensionnement du cluster s'exécute indéfiniment en raison de l'échec d'une seule opération de redimensionnement. Cela améliore également l'utilisation du cluster, car celui-ci augmente et diminue la capacité plus rapidement.
- Résout un problème selon lequel les opérations de réduction du cluster pouvaient être bloquées alors qu'Amazon EMR mettait hors service gracieusement un nœud principal et que celui-ci devenait inutilisable avant sa mise hors service complète.
- Améliore la stabilité d'un nœud dans un cluster à haute disponibilité comportant plusieurs nœuds principaux lorsqu'Amazon EMR redémarre un seul nœud.
- Optimise la gestion des journaux avec Amazon EMR exécuté sur Amazon. EC2 C'est pourquoi vous constaterez peut-être une légère réduction des coûts de stockage pour les journaux de votre cluster.
- Améliore la gestion des fichiers journaux de ZooKeeper transactions conservés sur les nœuds principaux afin de minimiser les scénarios dans lesquels les fichiers journaux dépassent les limites et interrompent les opérations du cluster.
- Corrige un bogue rare qui peut provoquer l'échec d'un cluster à haute disponibilité comportant plusieurs nœuds principaux en raison de l'impossibilité de communiquer avec le Yarn ResourceManager.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Taipei), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Pacifique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Malaisie), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)
2,0.2024 503.0	4,14,343	xxxxxx, 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

Versions des composants 5.36.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.16.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.5.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.21.0	Application de copie distribuée optimisée pour Amazon S3.

Composant	Version	Description
emr-s3-select	1.7.0	Connecteur S3Select EMR
emrfs	2,51,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.14.2	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.14.2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.10.1-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.10.1-amzn-4	HDFS node-level service for storing blocks.

Composant	Version	Description
hadoop-hdfs-library	2.10.1-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.10.1-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.10.1-amzn-4	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.10.1-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.10.1-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.10.1-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.10.1-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.10.1-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.13	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.9-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.9-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.9-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.9-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.9-amzn-2	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.9-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.9-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.10.1-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-spark	0.10.1-amzn-1	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hudi-presto	0.10.1-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.10.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.13.1	nginx [engine x] est un serveur HTTP et à proxy inverse

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.8.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68	Serveur de base de données MySQL.
nvidia-cuda	11,0,194	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.267-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.267-amzn-1	Service pour exécuter les éléments d'une requête.
presto-client	0.267-amzn-1	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.8-amzn-2	Clients de ligne de commande Spark.
spark-history-server	2.4.8-amzn-2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.8-amzn-2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.8-amzn-2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.10.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

5.36.2 classifications de configuration

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

classifications EMR-5.36.2

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Classifications	Description	Actions de reconfiguration
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Not available.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	Should not be reconfigured.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classifications	Description	Actions de reconfiguration
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.

Classifications	Description	Actions de reconfiguration
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.

Classifications	Description	Actions de reconfiguration
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.

Classifications	Description	Actions de reconfiguration
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.	Restarts EMR record server.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.	Restarts EMR record server.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.

Classifications	Description	Actions de reconfiguration
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.

Classifications	Description	Actions de reconfiguration
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

5.36.2 journal des modifications

Journal des modifications pour la version 5.36.2 et notes de version

Date	Événement	Description
28/05/2024	Déploiement terminé	Amazon EMR 5.36.2 entièrement déployé dans toutes les régions prises en charge
28/05/2024	Publication de documents	Notes de mise à jour d'Amazon EMR 5.36.2 publiées pour la première fois
20/05/2024	Première version	Amazon EMR 5.36.2 a été déployé pour la première fois dans des régions commercia les limitées

Amazon EMR version 5.36.1

Versions d’application 5.36.1

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, Iceberg, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
AWS SDK pour Java	1,12,206	1,12,206	1,12,159	1,1970
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.2	1.13.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
HCatalog	2.3.9	2.3.9	2.3.9	2.3.8
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.9	2.3.8
Hudi	0.10.1-amzn-1	0.10.1-amzn-1	0.9.0-amzn-2	0.9.0-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.6.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,267	0,267	0,266	0,261
Spark	2.4.8	2.4.8	2.4.8	2.4.8
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.10.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.36.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.36.1. Les modifications ont été apportées à la version 5.36.0. Pour plus d'informations sur le calendrier de publication, consultez le journal des modifications.

Modifications, améliorations et problèmes résolus

- La version 5.36.1 d'Amazon EMR ajoute la prise en charge de l'archivage des journaux dans Amazon S3 lors de la réduction de la taille du cluster. Dans les versions 5.x précédentes, vous pouviez uniquement archiver les fichiers journaux sur Amazon S3 lors de la résiliation du cluster. Cette amélioration garantit que les fichiers journaux générés sur le cluster sont conservés sur Amazon S3 même après la résiliation du nœud. Pour plus d'informations, consultez Configuration de la journalisation et du débogage de cluster.
- La version 5.36.1 améliore le démon de gestion des journaux sur le cluster afin de surveiller des dossiers de journaux supplémentaires dans votre cluster EMR. Cette amélioration permet de minimiser les scénarios de surutilisation des disques.
- La version 5.36.1 redémarre automatiquement le démon de gestion des journaux du cluster lorsqu'il s'arrête. Cette amélioration réduit le risque que les nœuds apparaissent défectueux en raison d'une surutilisation du disque.
- La version 5.36.1 corrige un problème selon lequel les démons Amazon EMR sur le nœud primaire conservaient des métadonnées périmées pour les instances résiliées du cluster. La conservation de données périmées peut entraîner une augmentation illimitée de l'utilisation du processeur et de la mémoire au sein du cluster et, en fin de compte, des défaillances du cluster.
- Pour les clusters lancés avec plusieurs nœuds principaux, la version 5.36.1 résout un problème selon lequel une défaillance EC2 matérielle d'Amazon sur l'un des nœuds principaux pouvait provoquer la défaillance d'un deuxième nœud principal et rendre votre cluster instable.

- Pour les clusters configurés avec un chiffrement en transit, Managed Scaling prend désormais en charge les données de réorganisation Spark. Les données de réorganisation Spark sont des données que Spark redistribue entre les partitions pour effectuer des opérations spécifiques. Lors de la réduction, Managed Scaling ignore les instances contenant des données de réorganisation. Cela permet d'éviter les nouvelles tentatives de tâches et les nouveaux calculs, qui sont coûteux en termes de prix et de performances. Pour plus d'informations sur les opérations de réorganisation, consultez le Guide de programmation Spark.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 822.0	4,1,4322	30 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 404.1	4,14.311	18 avril 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kon g), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

Versions des composants 5.36.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon

EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.16.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.5.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.21.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.7.0	Connecteur S3Select EMR
emrfs	2,51,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.

Composant	Version	Description
flink-client	1.14.2	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.14.2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.10.1-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.10.1-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.10.1-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.10.1-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	2.10.1-amzn-4	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.10.1-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.10.1-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.10.1-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.10.1-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.10.1-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.9-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.9-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.9-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.9-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.9-amzn-2	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.9-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.9-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.10.1-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-spark	0.10.1-amzn-1	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hudi-presto	0.10.1-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.10.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.13.1	nginx [engine x] est un serveur HTTP et à proxy inverse

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.8.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68	Serveur de base de données MySQL.
nvidia-cuda	11,0,194	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.267-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.267-amzn-1	Service pour exécuter les éléments d'une requête.
presto-client	0.267-amzn-1	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.8-amzn-2	Clients de ligne de commande Spark.
spark-history-server	2.4.8-amzn-2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.8-amzn-2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.8-amzn-2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.10.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.36.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-5.36.1

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Classifications	Description	Actions de reconfiguration
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Not available.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	Should not be reconfigured.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classifications	Description	Actions de reconfiguration
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.

Classifications	Description	Actions de reconfiguration
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
hudi-defaults	Modifiez les valeurs dans le fichier hudi-defaults.conf de Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.

Classifications	Description	Actions de reconfiguration
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.

Classifications	Description	Actions de reconfiguration
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.	Restarts EMR record server.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.	Restarts EMR record server.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.

Classifications	Description	Actions de reconfiguration
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.

Classifications	Description	Actions de reconfiguration
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

Journal des modifications 5.36.1

Journal des modifications pour la version 5.36.1 et les notes de version

Date	Événement	Description
26/07/2023	Mettre à jour	Étiquettes `2.0.20230612.0` et `2.0.20230628.0` de la nouvelle version du système d'exploitation.
25/05/2023	Déploiement terminé	Amazon EMR 5.36.1 entièrement déployé dans toutes les régions prises en charge
09/05/2023	Publication de documents	Notes de mise à jour d'Amazon EMR 5.36.1 publiées pour la première fois
04/05/2023	Première version	Amazon EMR 5.36.1 a été déployé pour la première fois dans des régions commercia les limitées

Amazon EMR version 5.36.0

Versions d’application 5.36.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, Iceberg, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
AWS SDK pour Java	1,12,206	1,12,159	1,1970	1,1970
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.13.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
HCatalog	2.3.9	2.3.9	2.3.8	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.8	2.3.7
Hudi	0.10.1-amzn-1	0.9.0-amzn-2	0.9.0-amzn-0	0.7.0-amzn-1
Hue	4.10.0	4.10.0	4.9.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.2.2
Livy	0.7.1	0.7.1	0.7.1	0.7.0
MXNet	1.8.0	1.8.0	1.8.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,267	0,266	0,261	0,245,1
Spark	2.4.8	2.4.8	2.4.8	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.36.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.36.0. Les modifications ont été apportées à la version 5.35.0.

Date de parution initiale : 15 juin 2022

Nouvelles fonctions

- La version 5.36.0 d'Amazon EMR ajoute la prise en charge du langage de définition de données (DDL) avec Apache Spark sur les clusters compatibles avec Apache Ranger. Cela vous permet d'utiliser Apache Ranger pour gérer l'accès à des opérations telles que la création, la modification et la suppression de bases de données et de tables d'un cluster Amazon EMR.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

> **Note**
>
> Cette version ne bénéficie plus de mises à jour automatiques de l'AMI puisqu'elle a été suivie d'une version supplémentaire de correctifs. La version du correctif est indiquée par le numéro qui suit la deuxième décimale (`6.8.1`). Pour savoir si vous utilisez la dernière version du correctif, consultez les versions disponibles dans le Guide des versions, ou consultez le menu déroulant des versions d'Amazon EMR lorsque vous créez un cluster dans la console, ou utilisez l'action d'API `ListReleaseLabels` ou de CLI `list-release-labels`. Pour être tenu au courant des nouvelles versions, abonnez-vous au flux RSS sur la page Quoi de neuf ?

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Paci fique (Jakarta), États-Uni s Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hyderabad), USA Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), USA Ouest (Oregon), USA Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Israël (Tel Aviv), Canada (Centre), Europe (Espagne), Europe (Zurich)
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			du Nord), USA Ouest (Oregon), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourne), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapou r), Asie-Pacifique (Sydney), Asie-Pacifique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfort), Europe (Irlande), Europe (Londres) , Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapou r) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfort), Europe (Irlande), Europe (Londres) , Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapou r) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfort), Europe (Irlande), Europe (Londres) , Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapou r) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfort), Europe (Irlande), Europe (Londres) , Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapou r) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfort), Europe (Irlande), Europe (Londres) , Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapou r) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfort), Europe (Irlande), Europe (Londres) , Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapou r) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfort), Europe (Irlande), Europe (Londres) , Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapou r) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfort), Europe (Irlande), Europe (Londres) , Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapou r) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe (Zurich), Asie-Pacifique

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Melbourne), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfort), Europe (Irlande), Europe (Londres) , Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapou r) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe (Zurich), Asie-Pacifique

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Melbourne), Israël (Tel Aviv)
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres) , Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 418,0	4,14.311	3 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres) , Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 404.1	4,14.311	18 avril 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres) , Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0.2023 404.0	4,14.311	10 avril 2023	USA Est (Virginie du Nord), Europe (Paris)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 320,0	4,14,309	30 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres) , Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 307.0	4,14,305	15 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres) , Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 207.0	4,14,304	22 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres) , Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 119.1	4,14,301	3 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres) , Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2022 210.1	4,14,301	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres) , Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 103.3	4,14.296	5 décembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres) , Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 004.0	4,14.294	2 novembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres) , Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 912.1	4,14,291	7 octobre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres) , Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0,2022 719,0	4,14.287	10 août 2022	USA Ouest (Californie du Nord), Europe (Paris), Europe (Stockholm), Europe (Francfort), Asie-Pacifique (Mumbai), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 426.0	4,14,281	14 juin 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres) , Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

Modifications, améliorations et problèmes résolus

- Les mises à niveau vers Amazon EMR 5.36.0 sont désormais prises en charge : aws-sdk 1.12.206, Hadoop 2.10.1-amzn-4, Hive 2.3.9-amzn-2, Hudi 0.10.1-amzn-1, Spark 2.4.8-amzn-2, Presto 0.267-amzn-1, connecteur Amazon Glue 1.18.0, EMRFS 2.51.0.

Problèmes connus

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.
- Avec les versions 5.36.0 et 6.6.0 à 6.9.0 d'Amazon EMR, les composants de service `SecretAgent` et `RecordServer` peuvent subir une perte de données de journal en raison d'une configuration incorrecte du modèle de nom de fichier dans les propriétés de Log4j2. En cas de configuration incorrecte, les composants ne génèrent qu'un seul fichier journal par jour. Lorsque la stratégie de rotation est appliquée, elle remplace le fichier existant au lieu de générer un nouveau fichier journal comme prévu. Pour contourner le problème, utilisez une action d'amorçage pour générer des journaux toutes les heures et ajoutez un nombre entier auto-incrémenté dans le nom du fichier pour gérer la rotation.

  Pour les versions 6.6.0 à 6.9.0 d'Amazon EMR, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
  ```

  Pour Amazon EMR 5.36.0, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Versions des composants 5.36.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.16.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.5.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.21.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.7.0	Connecteur S3Select EMR
emrfs	2,51,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.14.2	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.14.2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.10.1-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.10.1-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.10.1-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.10.1-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.10.1-amzn-4	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.10.1-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.10.1-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.10.1-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.10.1-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.10.1-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.9-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.3.9-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.9-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.9-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.9-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.9-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.9-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.10.1-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-spark	0.10.1-amzn-1	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hudi-presto	0.10.1-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.

Composant	Version	Description
hue-server	4.10.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.13.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.8.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68	Serveur de base de données MySQL.
nvidia-cuda	11,0,194	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur

Composant	Version	Description
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.267-amzn-1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.267-amzn-1	Service pour exécuter les éléments d'une requête.
presto-client	0.267-amzn-1	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.8-amzn-2	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.4.8-amzn-2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.8-amzn-2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.8-amzn-2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.10.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.36.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-5.36.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager,

Classifications	Description	Actions de reconfiguration
		NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Not available.

Classifications	Description	Actions de reconfiguration
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	Should not be reconfigured.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classifications	Description	Actions de reconfiguration
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.

Classifications	Description	Actions de reconfiguration
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.

Classifications	Description	Actions de reconfiguration
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.	Restarts EMR record server.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.	Restarts EMR record server.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environ nement Sqoop.	Not available.

Classifications	Description	Actions de reconfiguration
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.

Classifications	Description	Actions de reconfiguration
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Amazon EMR version 5.35.0

Versions d'application 5.35.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, Iceberg, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
AWS SDK pour Java	1,12,159	1,1970	1,1970	1,1970
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
Delta	-	-	-	-
Flink	1.14.2	1.13.1	1.12.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.8	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.8	2.3.7	2.3.7
Hudi	0.9.0-amzn-2	0.9.0-amzn-0	0.7.0-amzn-1	0.7.0-amzn-1
Hue	4.10.0	4.9.0	4.9.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.2.2	1.2.2
Livy	0.7.1	0.7.1	0.7.0	0.7.0
MXNet	1.8.0	1.8.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,266	0,261	0,245,1	0,245,1
Spark	2.4.8	2.4.8	2.4.7	2.4.7

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.35.0

Ceci est la note de mise à jour de la version 5.35.0 d'Amazon EMR.

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.35.0. Les modifications ont été apportées à la version 5.34.0.

Date de parution initiale : 30 mars 2022

Nouvelles fonctions

- Les applications Amazon EMR version 5.35 qui utilisent Log4j 1.x et Log4j 2.x sont mises à niveau pour utiliser Log4j 1.2.17 (ou version ultérieure) et Log4j 2.17.1 (ou version ultérieure) respectivement, et ne nécessitent pas d'actions d'amorçage pour atténuer les problèmes CVE dans les versions précédentes. Consultez Approche visant à atténuer le CVE-2021-44228.

Modifications, améliorations et problèmes résolus

Modifications apportées à Flink

Type de modification	Description
Mises à niveau	• Mise à jour de la version de flink vers la version 1.14.2. • log4j a été mis à jour vers la version 2.17.1.

Modifications apportées à Hadoop

Type de modification	Description
Rétroportages de Hadoop open source depuis EMR 5.34.0	• YARN-10438 : Gérer un identifiant de conteneur nul dans le numéro du client () RMService getContainerReport • YARN-7266 : les threads du gestionnaire d'événements du serveur de chronologie sont verrouillés • YARN-10438 : ATS 1.5 ne démarre pas si les RollingLevelDb fichiers sont corrompus ou manquants • HADOOP-13500 : Synchronisation de l'itérati on de l'objet Propriétés de configuration • YARN-10651 : CapacityScheduler s'est écrasé avec NPE dedans. AbstractY arnScheduler updateNodeResource() • HDFS-12221 : Remplacer les caractères dans XmlEditsVisitor • HDFS-16410 : Analyse XML non sécurisée OfflineEditsXmlLoader
Modifications et corrections apportées à Hadoop	• Tomcat utilisé dans KMS et HttpFS est mis à niveau vers la version 8.5.75 • Dans la FileSystemOptimizedCommitter version 2, le marqueur de réussite était écrit dans le chemin de sortie CommitJob défini lors de la création du commit. Étant donné que les chemins de sortie de CommitJob et de la tâche peuvent différer, le chemin a été corrigé pour utiliser celui défini dans les fichiers manifestes. Pour les tâches Hive, cela signifie que le marqueur de réussite est correctement écrit lors de l'exécution

Type de modification	Description
	d'opérations telles que la partition dynamique ou UNION ALL.

Modifications apportées à Hive

Type de modification	Description
Hive est passé à la version open source 2.3.9, y compris ces correctifs JIRA	• HIVE-17155 : findConfFile () dans le HiveConf fichier .java présente quelques problèmes avec le chemin de configuration • HIVE-24797 : désactivation de la validation des valeurs par défaut lors de l'analyse des schémas Avro • HIVE-21563 : Améliorez les performances de Table# getEmptyTable en le désactivant Once registerAllFunctions • HIVE-18147 : Les tests peuvent échouer avec java.net. BindException: Adresse déjà utilisée • HIVE-24608 : revenir à get_table dans le client HMS pour Hive 2.3.x • HIVE-21200 : Vectorisation : colonne de date lançant java.lang. UnsupportedOperati onException pour parquet • HIVE-19228 : suppression de l'utilisation de commons-httpclient 3.x
Rétroportages de Hive open source depuis EMR 5.34.0	• HIVE-19990 : la requête avec un intervalle littéral dans la condition de jointure échoue • HIVE-25824 : mise à niveau de la branche-2 .3 vers log4j 2.17.0

Type de modification	Description
	• TEZ-4062 : la planification des tentatives spéculatives doit être abandonnée une fois la tâche terminée • TEZ-4108 : NullPointerException lors d'une exécution spéculative, condition de course • TEZ-3918 : le paramètre tez.task.log.level ne fonctionne pas
Mises à niveau et correctifs apportés à Hive	• Mise à niveau de la version de Log4j vers 2.17.1 • Mise à niveau de la version de ORC vers 1.4.3 • Correction d'un blocage dû à un fil de pénalité ShuffleScheduler
Nouvelles fonctionnalités	• Ajout d'une fonctionnalité permettant d'imprimer Hive Query dans les journaux AM. Par défaut, l'option est désactivée. Drapeau/Conf :.`tez.am.emr.print.hive.query.in.log` Statut (par défaut) : FALSE.

Modifications apportées à Oozie

Type de modification	Description
Rétroportages d'Oozie open source depuis EMR 5.34.0	• OOZIE-3652 : Le lanceur Oozie devrait réessayer de répertorier les répertoires lorsque cela se produit NoSuchFileException

Modifications apportées à Pig

Type de modification	Description
Mises à niveau	• log4j a été mis à jour vers la version 1.2.17.

Problèmes connus

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&‘()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 5.35.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.15.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.5.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.20.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.7.0	Connecteur S3Select EMR
emrfs	2,49,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.14.2	Applications et scripts client de la ligne de commande Apache Flink.

Composant	Version	Description
flink-jobmanager-config	1.14.2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.10.1-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.10.1-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.10.1-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.10.1-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.10.1-amzn-3	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.

Composant	Version	Description
hadoop-httpfs-server	2.10.1-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.10.1-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.10.1-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.10.1-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.10.1-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.9-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.9-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.9-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.9-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.9-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.9-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.9-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.

Composant	Version	Description
hudi	0.9.0-amzn-2	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-spark	0.9.0-amzn-2	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hudi-presto	0.9.0-amzn-2	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.10.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.13.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.8.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68	Serveur de base de données MySQL.

Composant	Version	Description
nvidia-cuda	10,1243	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.266-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.266-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.266-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.

Composant	Version	Description
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.8-amzn-1	Clients de ligne de commande Spark.
spark-history-server	2.4.8-amzn-1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.8-amzn-1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.8-amzn-1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.10.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.35.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-5.35.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Not available.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	Should not be reconfigured.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classifications	Description	Actions de reconfiguration
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.

Classifications	Description	Actions de reconfiguration
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.

Classifications	Description	Actions de reconfiguration
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.	Restarts EMR record server.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.	Restarts EMR record server.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environ nement Sqoop.	Not available.

Classifications	Description	Actions de reconfiguration
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.

Classifications	Description	Actions de reconfiguration
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Amazon EMR version 5.34.0

Versions d'application 5.34.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Mahout, Oozie, PhoenixPig, Presto, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
AWS SDK pour Java	1,1970	1,1970	1,1970	1,1,890
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
Delta	-	-	-	-
Flink	1.13.1	1.12.1	1.12.1	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.8	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.8	2.3.7	2.3.7	2.3.7
Hudi	0.9.0-amzn-0	0.7.0-amzn-1	0.7.0-amzn-1	0.6.0-amzn-0
Hue	4.9.0	4.9.0	4.9.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.2.2	1.2.2	1.1.0
Livy	0.7.1	0.7.0	0.7.0	0.7.0
MXNet	1.8.0	1.7.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,261	0,245,1	0,245,1	0,240,1
Spark	2.4.8	2.4.7	2.4.7	2.4.7

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.9.0	0.9.0	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.34.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.34.0. Les modifications ont été apportées à la version 5.33.1.

Date de parution initiale : 20 janvier 2022

Date de mise à niveau : 21 mars 2022

Nouvelles fonctions

- [Mise à l'échelle gérée] Optimisation de la mise à l'échelle gérée des données de réorganisation Spark – Pour Amazon EMR versions 5.34.0 et ultérieures, et EMR versions 6.4.0 et ultérieures, la mise à l'échelle gérée prend désormais en compte les données de réorganisation Spark (données que Spark redistribue entre les partitions pour effectuer des opérations spécifiques). Pour plus d'informations sur les opérations de réorganisation, consultez Utilisation de la mise à l'échelle gérée par EMR dans Amazon EMR dans le Guide de gestion Amazon EMR et le Guide de programmation Spark.
- [Hudi] Améliorations visant à simplifier la configuration de Hudi. Le contrôle de simultanéité optimiste a été désactivé par défaut.

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.
- Auparavant, le redémarrage manuel du gestionnaire de ressources sur un cluster multimaître provoquait le rechargement par les démons Amazon EMR on-cluster, comme Zookeeper, de tous les nœuds précédemment mis hors service ou perdus dans le fichier znode de Zookeeper. Cela a entraîné le dépassement des limites par défaut dans certaines situations. Amazon EMR supprime désormais les enregistrements de nœuds mis hors service ou perdus datant de plus d'une heure du fichier Zookeeper et les limites internes ont été augmentées.
- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Mise à niveau de Zeppelin vers la version 0.10.0.
- Livy Fix – mise à niveau vers la version 0.7.1
- Amélioration des performances de Spark – les exécuteurs hétérogènes sont désactivés lorsque certaines valeurs de configuration de Spark sont remplacées dans EMR 5.34.0.

- WebHDFS et le serveur HttpFS sont désactivés par défaut. Vous pouvez réactiver WebHDFS en utilisant la configuration Hadoop, `dfs.webhdfs.enabled`. Le serveur HttpFS peut être démarré en utilisant `sudo systemctl start hadoop-httpfs`.

Problèmes connus

- La fonctionnalité Blocs-notes Amazon EMR utilisée avec l'emprunt d'identité de l'utilisateur Livy ne fonctionne pas car HttpFS est désactivé par défaut. Dans ce cas, le bloc-notes EMR ne peut pas se connecter au cluster dont l'emprunt d'identité Livy est activé. La solution consiste à démarrer le serveur HttpFS avant de connecter le bloc-notes EMR au cluster à l'aide de `sudo systemctl start hadoop-httpfs`.
- Les requêtes Hue ne fonctionnent pas dans Amazon EMR 6.4.0 car le serveur Apache Hadoop HTTPFS est désactivé par défaut. Pour utiliser Hue sur Amazon EMR 6.4.0, démarrez manuellement le serveur HttpFS sur le nœud primaire d'Amazon EMR à l'aide de `sudo systemctl start hadoop-httpfs`, ou utilisez une étape d'Amazon EMR.
- La fonctionnalité Blocs-notes Amazon EMR utilisée avec l'emprunt d'identité de l'utilisateur Livy ne fonctionne pas car HttpFS est désactivé par défaut. Dans ce cas, le bloc-notes EMR ne peut pas se connecter au cluster dont l'emprunt d'identité Livy est activé. La solution consiste à démarrer le serveur HttpFS avant de connecter le bloc-notes EMR au cluster à l'aide de `sudo systemctl start hadoop-httpfs`.
- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 5.34.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.14.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-notebook-env	1.4.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.18.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.7.0	Connecteur S3Select EMR
emrfs	2,48,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.13.1	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.13.1	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.10.1-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.10.1-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.10.1-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.10.1-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.10.1-amzn-2	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.10.1-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.10.1-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.10.1-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.10.1-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.10.1-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.8-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.3.8-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.8-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.8-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.8-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.8-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.8-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.9.0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-spark	0.9.0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hudi-presto	0.9.0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.

Composant	Version	Description
hue-server	4.9.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	1.4.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.1-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.13.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.8.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68	Serveur de base de données MySQL.
nvidia-cuda	10,1243	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.1	Client de ligne de commande Oozie.
oozie-server	5.2.1	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur

Composant	Version	Description
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.261-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.261-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.261-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.8-amzn-0	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.4.8-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.8-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.8-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.10.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.34.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-5.34.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager,

Classifications	Description	Actions de reconfiguration
		NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Not available.

Classifications	Description	Actions de reconfiguration
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	Should not be reconfigured.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classifications	Description	Actions de reconfiguration
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.

Classifications	Description	Actions de reconfiguration
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.

Classifications	Description	Actions de reconfiguration
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.	Restarts EMR record server.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.	Restarts EMR record server.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environ nement Sqoop.	Not available.

Classifications	Description	Actions de reconfiguration
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.

Classifications	Description	Actions de reconfiguration
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Amazon EMR version 5.33.1

Versions d'application 5.33.1

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Mahout, Oozie, PhoenixPig, Presto, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
AWS SDK pour Java	1,1970	1,1970	1,1,890	1,1,890
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
Delta	-	-	-	-
Flink	1.12.1	1.12.1	1.11.2	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.7.0-amzn-1	0.7.0-amzn-1	0.6.0-amzn-0	0.6.0-amzn-0
Hue	4.9.0	4.9.0	4.8.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.2.2	1.2.2	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,245,1	0,245,1	0,240,1	0,240,1
Spark	2.4.7	2.4.7	2.4.7	2.4.7

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.3.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.9.0	0.9.0	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.33.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.33.0/5.33.1. Les modifications ont été apportées à la version 5.32.0.

Date de parution initiale : 19 avril 2021

Dernière mise à jour : 9 août 2021

Mises à niveau

- Mise à niveau du connecteur Amazon Glue vers la version 1.15.0
- Mise à niveau AWS SDK pour Java vers la version 1.11.970
- Mise à niveau de EMRFS vers la version 2.46.0
- Mise à niveau de EMR Goodies vers la version 2.14.0
- Mise à niveau du serveur d'enregistrement EMR vers la version 1.9.0
- Mise à niveau de EMR S3 Dist CP vers la version 2.18.0
- Mise à niveau de EMR Secret Agent vers la version 1.8.0
- Mise à niveau de Flink vers la version 1.12.1
- Mise à niveau de Hadoop vers la version 2.10.1-amzn-1
- Mise à niveau de Hive vers la version 2.3.7-amzn-4
- Mise à niveau de Hudi vers la version 0.7.0

- Mise à niveau de Hue vers la version 4.9.0
- Mise à niveau de OpenCV vers la version 4.5.0
- Mise à niveau de Presto vers la version 0.245.1-amzn-0
- Mise à niveau de R vers la version 4.0.2
- Mise à niveau de Spark vers la version 2.4.7-amzn-1
- Mise à niveau TensorFlow vers la version 2.4.1
- Mise à niveau de Zeppelin vers la version 0.9.0

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.
- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions

5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.

- Configuration d'un cluster pour résoudre les problèmes de performances d'Apache YARN Timeline Server versions 1 et 1.5

  Les versions 1 et 1.5 d'Apache YARN Timeline Server peuvent entraîner des problèmes de performances avec de grands clusters EMR très actifs, en particulier avec `yarn.resourcemanager.system-metrics-publisher.enabled=true`, le paramètre par défaut d'Amazon EMR. YARN Timeline Server v2 open source résout le problème de performance lié à la capacité de mise à l'échelle de YARN Timeline Server.

  Les autres solutions à ce problème incluent :

  - Configuration de yarn.resourcemanager. system-metrics-publisher.enabled=false dans le fichier yarn-site.xml.
  - Activation du correctif pour ce problème lors de la création d'un cluster, comme décrit ci-dessous.

  Les versions Amazon EMR suivantes contiennent un correctif pour ce problème de performance de YARN Timeline Server.

  EMR 5.30.2, 5.31.1, 5.32.1, 5.33.1, 5.34.x, 6.0.1, 6.1.1, 6.2.1, 6.3.1, 6.4.x

  Pour activer le correctif sur l'une des versions Amazon EMR spécifiées ci-dessus, définissez ces propriétés sur `true` dans un fichier JSON de configuration transmis à l'aide du paramètre de commande `aws emr create-cluster` : `--configurations file://./configurations.json`. Vous pouvez également activer le correctif à l'aide de l'interface utilisateur de la console de reconfiguration.

  Exemple du contenu du fichier configurations.json :

```
[
{
"Classification": "yarn-site",
"Properties": {
"yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
 "true",
"yarn.resourcemanager.system-metrics-publisher.enabled": "true"
},
```

```
"Configurations": []
}
]
```

- L'exécution de Spark est maintenant plus rapide lors de la récupération des emplacements de partition à partir de Hive Metastore pour les requêtes d'insertion de Spark.
- Versions de composants mises à niveau. Pour obtenir la liste des versions des composants, consultez la section À propos des versions d'Amazon EMR dans ce guide.
- Le bundle AWS Java SDK a été installé sur chaque nouveau cluster. Il s'agit d'un fichier jar unique contenant tous les services SDKs et leurs dépendances, au lieu de fichiers jar de composants individuels. Pour plus d'informations, consultez Dépendance du kit SDK Java groupé.
- Correction des problèmes de Managed Scaling dans les versions précédentes d'Amazon EMR et améliorations permettant de réduire considérablement les taux d'échec des applications.
- Le protocole HTTPS est désormais activé par défaut pour les référentiels Amazon Linux. Si vous utilisez une politique VPCE Amazon S3 pour restreindre l'accès à des compartiments spécifiques, vous devez ajouter le nouvel ARN du compartiment Amazon Linux `arn:aws:s3:::amazonlinux-2-repos-$region/*` à votre politique (remplacez `$region` par la région où se trouve le point de terminaison). Pour plus d'informations, consultez cette rubrique dans les forums de AWS discussion. Annonce : Amazon Linux 2 permet désormais d'utiliser le protocole HTTPS lors de la connexion aux référentiels de packages.

Nouvelles fonctionnalités

- Amazon EMR prend en charge les points d'accès Amazon S3, une fonctionnalité d'Amazon S3 qui vous permet de gérer facilement l'accès aux lacs de données partagés. En utilisant votre alias de point d'accès Amazon S3, vous pouvez simplifier l'accès aux données à grande échelle sur Amazon EMR. Vous pouvez utiliser les points d'accès Amazon S3 avec toutes les versions d'Amazon EMR sans frais supplémentaires dans toutes les AWS régions où Amazon EMR est disponible. Pour en savoir plus sur les points d'accès Amazon S3 et les alias de points d'accès, consultez Utilisation d'un alias de type compartiment pour votre point d'accès dans le Guide de l'utilisateur Amazon S3.
- Amazon EMR-5.33 prend en charge les nouveaux types d' EC2 instances Amazon : c5a, c5ad, c6gn, c6gd, m6gd, d3, d3en, m5zn, r5b, r6gd. Consultez Types d'instances pris en charge.

Problèmes connus

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Redémarrer InstanceController

     ```
     $ sudo systemctl daemon-reload

     $ sudo systemctl restart instance-controller
     ```

  Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Pour les clusters de sous-réseaux privés Amazon EMR 6.3.0 et 6.2.0, vous ne pouvez pas accéder à l'interface utilisateur Web de Ganglia. Vous recevrez un message d'erreur « accès refusé (403) ». D'autres sites Web UIs, tels que Spark, Hue JupyterHub, Zeppelin, Livy et Tez, fonctionnent normalement. L'accès à l'interface utilisateur Web de Ganglia sur les clusters de sous-réseaux publics fonctionne également normalement. Pour résoudre ce problème, redémarrez le service httpd sur le nœud primaire avec `sudo systemctl restart httpd`. Ce problème est résolu dans Amazon EMR 6.4.0.
- **Important**

  Les clusters EMR qui exécutent Amazon Linux ou Amazon Linux 2 Amazon Machine Images (AMIs) utilisent le comportement par défaut d'Amazon Linux et ne téléchargent ni n'installent automatiquement les mises à jour importantes et critiques du noyau nécessitant un redémarrage. Ce comportement est identique à celui des autres EC2 instances Amazon qui exécutent l'AMI Amazon Linux par défaut. Si de nouvelles mises à jour logicielles Amazon Linux nécessitant un redémarrage (telles que les mises à jour du noyau, de NVIDIA et de CUDA) sont disponibles après la publication d'une version d'Amazon EMR, les instances de cluster EMR qui exécutent l'AMI par défaut ne téléchargent pas et

n'installent pas automatiquement ces mises à jour. Pour obtenir les mises à jour du noyau, vous pouvez personnaliser votre AMI Amazon EMR afin d'utiliser la dernière AMI Amazon Linux.

- La prise en charge par console pour créer une configuration de sécurité spécifiant l'option d'intégration de AWS Ranger n'est actuellement pas prise en charge dans la GovCloud région. La configuration de la sécurité peut être effectuée à l'aide de la CLI. Consultez Création de la configuration de sécurité EMR dans le Guide de gestion Amazon EMR.
- Politiques gérées délimitées : pour s'aligner sur les AWS meilleures pratiques, Amazon EMR a introduit des politiques gérées par défaut définies dans la version 2 EMR en remplacement des politiques qui seront déconseillées. Consultez Politiques gérées par Amazon EMR.
- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère / (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 5.33.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon

EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.14.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.2.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.18.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.6.0	Connecteur S3Select EMR
emrfs	2,46,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.

Composant	Version	Description
flink-client	1.12.1	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.12.1	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.10.1-amzn-1.1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.10.1-amzn-1.1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.10.1-amzn-1.1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.10.1-amzn-1.1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	2.10.1-amzn-1.1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.10.1-amzn-1.1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.10.1-amzn-1.1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-1.1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.10.1-amzn-1.1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.10.1-amzn-1.1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.10.1-amzn-1.1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.7-amzn-4	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.7-amzn-4	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.7-amzn-4	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.7-amzn-4	Client de ligne de commande Hive.
hive-hbase	2.3.7-amzn-4	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.7-amzn-4	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.7-amzn-4	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.7.0-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-spark	0.7.0-amzn-1	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hudi-presto	0.7.0-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.9.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.2.2	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.7.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 ans et plus	Serveur de base de données MySQL.
nvidia-cuda	10,1243	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.245.1-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.245.1-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.245.1-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.7-amzn-1.1	Clients de ligne de commande Spark.
spark-history-server	2.4.7-amzn-1.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.7-amzn-1.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.7-amzn-1.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.9.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.33.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-5.33.1

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Classifications	Description	Actions de reconfiguration
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Not available.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	Should not be reconfigured.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classifications	Description	Actions de reconfiguration
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.

Classifications	Description	Actions de reconfiguration
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.

Classifications	Description	Actions de reconfiguration
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.	Restarts EMR record server.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.	Restarts EMR record server.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environ nement Sqoop.	Not available.

Classifications	Description	Actions de reconfiguration
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètre s de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.

Classifications	Description	Actions de reconfiguration
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Amazon EMR version 5.33.0

Versions d'application 5.33.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Mahout, Oozie, PhoenixPig, Presto, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
AWS SDK pour Java	1,1970	1,1,890	1,1,890	1,1,852
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
Delta	-	-	-	-
Flink	1.12.1	1.11.2	1.11.2	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.0
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.7.0-amzn-1	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0
Hue	4.9.0	4.8.0	4.8.0	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	-
JupyterHub	1.2.2	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.6.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,245,1	0,240,1	0,240,1	0,238,3
Spark	2.4.7	2.4.7	2.4.7	2.4.6

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.3.1	2.3.1	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.9.0	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.33.0

Versions des composants 5.33.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark

Composant	Version	Description
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.14.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.2.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.18.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.6.0	Connecteur S3Select EMR
emrfs	2,46,0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.12.1	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.12.1	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.10.1-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.10.1-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.10.1-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.10.1-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.10.1-amzn-1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.10.1-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.10.1-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.10.1-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.10.1-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.10.1-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.10.1-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.7-amzn-4	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.7-amzn-4	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.7-amzn-4	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.7-amzn-4	Client de ligne de commande Hive.
hive-hbase	2.3.7-amzn-4	Client Hive-hbase.
hive-metastore-server	2.3.7-amzn-4	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.7-amzn-4	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.7.0-amzn-1	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.

Composant	Version	Description
hudi-spark	0.7.0-amzn-1	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hudi-presto	0.7.0-amzn-1	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.9.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	1.2.2	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.7.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68	Serveur de base de données MySQL.
nvidia-cuda	10,1243	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.5.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.245.1-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.245.1-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.245.1-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	4.0.2	Projet R pour les calculs statistiques

Composant	Version	Description
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.7-amzn-1	Clients de ligne de commande Spark.
spark-history-server	2.4.7-amzn-1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.7-amzn-1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.7-amzn-1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.4.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.9.0	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.33.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-5.33.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-session	Modifiez les paramètres Flink log4j-session.properties pour la session. Kubernetes/Yarn	Not available.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Classifications	Description	Actions de reconfiguration
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Classifications	Description	Actions de reconfiguration
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	Should not be reconfigured.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat Server.

Classifications	Description	Actions de reconfiguration
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.

Classifications	Description	Actions de reconfiguration
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Classifications	Description	Actions de reconfiguration
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server.

Classifications	Description	Actions de reconfiguration
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.	Restarts EMR record server.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.	Restarts EMR record server.

Classifications	Description	Actions de reconfiguration
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.	Not available.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restarts Oozie.

Classifications	Description	Actions de reconfiguration
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zeppelin-site	Modification des paramètres de configuration dans zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.	Restarts Zookeeper server.

Amazon EMR version 5.32.1

Versions d’application 5.32.1

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Mahout, Oozie, PhoenixPig, Presto, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
AWS SDK pour Java	1,1,890	1,1,890	1,1,852	1,1,852
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.2	1.11.0	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.0	2.10.0
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
Hue	4.8.0	4.8.0	4.7.1	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.6.0	1.6.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,240,1	0,240,1	0,238,3	0,238,3
Spark	2.4.7	2.4.7	2.4.6	2.4.6
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.3.1	2.1.0	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.32.1

Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.

Modifications, améliorations et problèmes résolus

- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Le protocole HTTPS est désormais activé par défaut pour les référentiels Amazon Linux. Si vous utilisez une politique VPCE Amazon S3 pour restreindre l'accès à des compartiments spécifiques, vous devez ajouter le nouvel ARN du compartiment Amazon Linux `arn:aws:s3:::amazonlinux-2-repos-$region/*` à votre politique (remplacez `$region` par

la région où se trouve le point de terminaison). Pour plus d'informations, consultez cette rubrique dans les forums de AWS discussion. Annonce : Amazon Linux 2 permet désormais d'utiliser le protocole HTTPS lors de la connexion aux référentiels de packages.

Problèmes connus

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 5.32.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*`-amzn-`*EmrVersion*. *EmrVersion*

commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.13.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.1.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.17.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.6.0	Connecteur S3Select EMR
emrfs	2,45,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.11.2	Applications et scripts client de la ligne de commande Apache Flink.

Composant	Version	Description
flink-jobmanager-config	1.11.2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.10.1-amzn-0.1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.10.1-amzn-0.1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.10.1-amzn-0.1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.10.1-amzn-0.1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.10.1-amzn-0.1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.

Composant	Version	Description
hadoop-httpfs-server	2.10.1-amzn-0.1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.10.1-amzn-0.1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-0.1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.10.1-amzn-0.1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.10.1-amzn-0.1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.10.1-amzn-0.1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.7-amzn-3	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.7-amzn-3	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.7-amzn-3	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.7-amzn-3	Client de ligne de commande Hive.
hive-hbase	2.3.7-amzn-3	Client Hive-hbase.
hive-metastore-server	2.3.7-amzn-3	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.7-amzn-3	Service pour l'acceptation de requêtes Hive en tant que requêtes web.

Composant	Version	Description
hudi	0.6.0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-spark	0.6.0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hudi-presto	0.6.0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.8.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	1.1.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.7.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68 et plus	Serveur de base de données MySQL.

Composant	Version	Description
nvidia-cuda	10,1243	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.240.1-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.240.1-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.240.1-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.

Composant	Version	Description
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.3	Projet R pour les calculs statistiques
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.7-amzn-0.1	Clients de ligne de commande Spark.
spark-history-server	2.4.7-amzn-0.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.7-amzn-0.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.7-amzn-0.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.3.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.8.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.32.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-5.32.1

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.

Classifications	Description	Actions de reconfiguration
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.

Classifications	Description	Actions de reconfiguration
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.	Not available.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	Should not be reconfigured.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classifications	Description	Actions de reconfiguration
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.

Classifications	Description	Actions de reconfiguration
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.

Classifications	Description	Actions de reconfiguration
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.	Restarts EMR record server.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.	Restarts EMR record server.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environ nement Sqoop.	Not available.

Classifications	Description	Actions de reconfiguration
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

Amazon EMR version 5.32.0

Versions d’application 5.32.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHubLivy, MXNet, Mahout, Oozie, PhoenixPig, Presto, Spark, Sqoop, TensorFlowTez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.32.0	emr-5.31.1	emr-5.31.0	emr-5.30.2
AWS SDK pour Java	1,1,890	1,1,852	1,1,852	1,1,759
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.0	1.11.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13

	emr-5.32.0	emr-5.31.1	emr-5.31.0	emr-5.30.2
HCatalog	2.3.7	2.3.7	2.3.7	2.3.6
Hadoop	2.10.1	2.10.0	2.10.0	2.8.5
Hive	2.3.7	2.3.7	2.3.7	2.3.6
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0	0.5.2-incubating
Hue	4.8.0	4.7.1	4.7.1	4.6.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.6.0	1.6.0	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,240,1	0,238,3	0,238,3	0.232
Spark	2.4.7	2.4.6	2.4.6	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.1.0	2.1.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.32.0	emr-5.31.1	emr-5.31.0	emr-5.30.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.32.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.32.0. Les modifications ont été apportées à la version 5.31.0.

Date de parution initiale : 8 janvier 2021

Mises à niveau

- Mise à niveau du connecteur Amazon Glue vers la version 1.14.0
- Mise à niveau du SDK Amazon SageMaker Spark vers la version 1.4.1
- Mise à niveau AWS SDK pour Java vers la version 1.11.890
- Mise à niveau de la version 4.16.0 du connecteur DynamoDB d'EMR
- Mise à niveau d'EMRFS vers la version 2.45.0
- Mise à niveau des métriques d'analyse des journaux d'EMR vers la version 1.18.0
- MetricsAndEventsApiGateway Client EMR mis à niveau vers la version 1.5.0
- Mise à niveau du serveur d'enregistrement EMR vers la version 1.8.0
- Mise à niveau d'EMR S3 Dist CP vers la version 2.17.0
- Mise à niveau d'EMR Secret Agent vers la version 1.7.0
- Mise à niveau de Flink vers la version 1.11.2
- Mise à niveau de Hadoop vers la version 2.10.1-amzn-0
- Mise à niveau de Hive vers la version 2.3.7-amzn-3
- Mise à niveau de Hue vers la version 4.8.0
- Mise à niveau de Mxnet vers la version 1.7.0

- Mise à niveau d'OpenCV vers la version 4.4.0
- Mise à niveau de Presto vers la version 0.240.1-amzn-0
- Mise à niveau de Spark vers la version 2.4.7-amzn-0
- Mise à niveau TensorFlow vers la version 2.3.1

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.
- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.

- Versions de composants mises à niveau.
- Pour obtenir la liste des versions des composants, consultez la section À propos des versions d'Amazon EMR dans ce guide.

Nouvelles fonctionnalités

- À partir d'Amazon EMR 5.32.0 et 6.5.0, le dimensionnement dynamique de l'exécuteur pour Apache Spark est activé par défaut. Pour activer ou désactiver cette fonctionnalité, vous pouvez utiliser le paramètre de configuration `spark.yarn.heterogeneousExecutors.enabled`.
- État de prise en charge du service de métadonnées d'instance (IMDS) V2 : les composants Amazon EMR 5.23.1, 5.27.1 et 5.32 ou versions ultérieures sont utilisés pour tous les appels IMDS. IMDSv2 Pour les appels IMDS dans le code de votre application, vous pouvez utiliser les deux IMDSv1 ou configurer l'IMDS pour qu'il ne soit utilisé que IMDSv2 pour renforcer la sécurité. IMDSv2 Pour les autres versions 5.x EMR, la IMDSv1 désactivation entraîne l'échec du démarrage du cluster.
- À partir d'Amazon EMR 5.32.0, vous pouvez lancer un cluster qui s'intègre nativement à Apache Ranger. Apache Ranger est un cadre open source permettant d'activer, de surveiller et de gérer la sécurité globale des données sur la plateforme Hadoop. Pour plus d'informations, consultez Apache Ranger. Grâce à l'intégration native, vous pouvez utiliser votre propre Apache Ranger pour appliquer un contrôle précis de l'accès aux données sur Amazon EMR. Consultez Intégration d'Amazon EMR à Apache Ranger dans le Guide de version Amazon EMR.
- La version 5.32.0 d'Amazon EMR prend en charge Amazon EMR sur EKS. Pour en savoir plus sur la prise en main d'EMR sur EKS, consultez Qu'est-ce qu'Amazon EMR sur EKS ?.
- La version 5.32.0 d'Amazon EMR prend en charge Amazon EMR Studio (version préliminaire). Pour plus d'informations sur la prise en main d'EMR Studio, consultez Amazon EMR Studio (version préliminaire).
- Politiques gérées délimitées : pour s'aligner sur les AWS meilleures pratiques, Amazon EMR a introduit des politiques gérées par défaut définies dans la version 2 EMR en remplacement des politiques qui seront déconseillées. Consultez Politiques gérées par Amazon EMR.

Problèmes connus

- Pour les clusters de sous-réseaux privés Amazon EMR 6.3.0 et 6.2.0, vous ne pouvez pas accéder à l'interface utilisateur Web de Ganglia. Vous recevrez un message d'erreur « accès refusé (403) ». D'autres sites Web UIs, tels que Spark, Hue JupyterHub, Zeppelin, Livy et Tez, fonctionnent

normalement. L'accès à l'interface utilisateur Web de Ganglia sur les clusters de sous-réseaux publics fonctionne également normalement. Pour résoudre ce problème, redémarrez le service httpd sur le nœud primaire avec `sudo systemctl restart httpd`. Ce problème est résolu dans Amazon EMR 6.4.0.

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Redémarrer InstanceController

     ```
     $ sudo systemctl daemon-reload

     $ sudo systemctl restart instance-controller
     ```

Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Les clusters EMR qui exécutent Amazon Linux ou Amazon Linux 2 Amazon Machine Images (AMIs) utilisent le comportement par défaut d'Amazon Linux et ne téléchargent ni n'installent automatiquement les mises à jour importantes et critiques du noyau nécessitant un redémarrage. Ce comportement est identique à celui des autres EC2 instances Amazon qui exécutent l'AMI Amazon Linux par défaut. Si de nouvelles mises à jour logicielles Amazon Linux nécessitant un redémarrage (telles que les mises à jour du noyau, de NVIDIA et de CUDA) sont disponibles après la publication d'une version d'Amazon EMR, les instances de cluster EMR qui exécutent l'AMI par défaut ne téléchargent pas et n'installent pas automatiquement ces mises à jour. Pour obtenir les mises à jour du noyau, vous pouvez personnaliser votre AMI Amazon EMR afin d'utiliser la dernière AMI Amazon Linux.

- La prise en charge par console pour créer une configuration de sécurité spécifiant l'option d'intégration de AWS Ranger n'est actuellement pas prise en charge dans la GovCloud région. La configuration de la sécurité peut être effectuée à l'aide de la CLI. Consultez Création de la configuration de sécurité EMR dans le Guide de gestion Amazon EMR.
- Lorsque AtRestEncryption le chiffrement HDFS est activé sur un cluster qui utilise Amazon EMR 5.31.0 ou 5.32.0, les requêtes Hive génèrent l'exception d'exécution suivante.

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
 attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
 java.lang.RuntimeException: Hive Runtime Error while closing
 operators: java.io.IOException: java.util.ServiceConfigurationError:
 org.apache.hadoop.security.token.TokenIdentifier: Provider
 org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère / (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 5.32.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont

propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,16,0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.13.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.1.0	Environnement Conda pour le bloc-notes EMR qui inclut la passerelle Jupyter Entreprise
emr-s3-dist-cp	2.17.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.6.0	Connecteur S3Select EMR

Composant	Version	Description
emrfs	2,45,0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.11.2	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.11.2	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.10.1-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.10.1-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.10.1-amzn-0	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	2.10.1-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.10.1-amzn-0	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.10.1-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.10.1-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.10.1-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.10.1-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.10.1-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.7-amzn-3	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.7-amzn-3	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.7-amzn-3	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.7-amzn-3	Client de ligne de commande Hive.
hive-hbase	2.3.7-amzn-3	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.7-amzn-3	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.7-amzn-3	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.6.0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-spark	0.6.0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hudi-presto	0.6.0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.8.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.1.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.7.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5,5,68	Serveur de base de données MySQL.
nvidia-cuda	10,1243	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.240.1-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.240.1-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.240.1-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.3	Projet R pour les calculs statistiques
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.7-amzn-0	Clients de ligne de commande Spark.
spark-history-server	2.4.7-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.7-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.7-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	2.3.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.32.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les actions de reconfiguration se produisent lorsque vous spécifiez une configuration pour les groupes d'instances d'un cluster en cours d'exécution. Amazon EMR lance uniquement des actions de reconfiguration pour les classifications que vous modifiez. Pour de plus amples informations, veuillez consulter Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Classifications emr-5.32.0

Classifications	Description	Actions de reconfiguration
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Modifiez les valeurs dans le fichier container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.	Not available.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Modifiez les paramètres liés au docker.	Not available.
emrfs-site	Modifiez les paramètres EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Classifications	Description	Actions de reconfiguration
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Modifiez les paramètres flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Modifiez les paramètres log4j.properties Flink.	Restarts Flink history server.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.	Not available.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.	Restarts Flink history server.

Classifications	Description	Actions de reconfiguration
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop	Not available.
hadoop-ssl-client	Modifier la configuration du client ssl hadoop	Not available.

Classifications	Description	Actions de reconfiguration
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Changez les valeurs dans HBase l'environnement.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-me trics2-hbase.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-pol icy.xml.	Not available.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-sit e.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.	Should not be reconfigured.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Classifications	Description	Actions de reconfiguration
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.	Not available.
hive-env	Modifiez les valeurs dans l'environnement Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.

Classifications	Description	Actions de reconfiguration
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.	Not available.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.	Not available.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2	Not available.
hue-ini	Modifiez les valeurs dans le fichier ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.	Not available.

Classifications	Description	Actions de reconfiguration
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.	Not available.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.	Not available.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.	Not available.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.	Restarts Livy Server.
livy-env	Modifiez les valeurs dans l'environnement Livy.	Restarts Livy Server.

Classifications	Description	Actions de reconfiguration
livy-log4j	Modifiez les paramètres Livy log4j.properties.	Restarts Livy Server.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.	Restarts Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.	Restarts Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hb ase.properties de Phoenix.	Not available.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.	Restarts Phoenix-QueryServe r.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-ph oenix.properties de Phoenix.	Not available.

Classifications	Description	Actions de reconfiguration
pig-env	Modifiez les valeurs dans l'environnement Pig.	Not available.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.	Restarts Oozie.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.	Not available.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.	Restarts Presto-Server.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authentic ator.properties de Presto.	Not available.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.	Not available.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.	Not available.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.	Not available.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.	Not available.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.	Not available.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.	Not available.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.	Not available.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.	Not available.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.	Not available.

Classifications	Description	Actions de reconfiguration
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.	Not available.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.	Not available.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.	Not available.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.	Not available.

Classifications	Description	Actions de reconfiguration
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.	Restarts EMR record server.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.	Restarts EMR record server.
spark	Paramètres définis par Amazon EMR pour Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Modifiez les valeurs dans l'environnement Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark	Not available.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Modifiez les valeurs d'environ nement Sqoop.	Not available.

Classifications	Description	Actions de reconfiguration
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.	Not available.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.	Not available.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Modifiez les valeurs dans l'environnement YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.	Restarts Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.pro perties.	Restarts Zookeeper server.

Amazon EMR version 5.31.1

Versions d’application 5.31.1

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.31.1	emr-5.31.0	emr-5.30.2	emr-5.30.1
AWS SDK pour Java	1,1,852	1,1,852	1,1,759	1,1,759
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	1.10.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13

	emr-5.31.1	emr-5.31.0	emr-5.30.2	emr-5.30.1
HCatalog	2.3.7	2.3.7	2.3.6	2.3.6
Hadoop	2.10.0	2.10.0	2.8.5	2.8.5
Hive	2.3.7	2.3.7	2.3.6	2.3.6
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.5.2-incubating	0.5.2-incubating
Hue	4.7.1	4.7.1	4.6.0	4.6.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,238,3	0,238,3	0.232	0.232
Spark	2.4.6	2.4.6	2.4.5	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.31.1	emr-5.31.0	emr-5.30.2	emr-5.30.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.31.1

Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.

Modifications, améliorations et problèmes résolus

- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.

- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Le protocole HTTPS est désormais activé par défaut pour les référentiels Amazon Linux. Si vous utilisez une politique VPCE Amazon S3 pour restreindre l'accès à des compartiments spécifiques, vous devez ajouter le nouvel ARN du compartiment Amazon Linux `arn:aws:s3:::amazonlinux-2-repos-$region/*` à votre politique (remplacez `$region` par la région où se trouve le point de terminaison). Pour plus d'informations, consultez cette rubrique dans les forums de AWS discussion. Annonce : Amazon Linux 2 permet désormais d'utiliser le protocole HTTPS lors de la connexion aux référentiels de packages.

Problèmes connus

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 5.31.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,15,0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.13.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.15.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.6.0	Connecteur S3Select EMR

Composant	Version	Description
emrfs	2.43.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.11.0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.11.0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.10.0-amzn-0.1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.10.0-amzn-0.1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.10.0-amzn-0.1	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	2.10.0-amzn-0.1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.10.0-amzn-0.1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.10.0-amzn-0.1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.10.0-amzn-0.1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.10.0-amzn-0.1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.10.0-amzn-0.1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.10.0-amzn-0.1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.10.0-amzn-0.1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.7-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.7-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.7-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.7-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.7-amzn-1	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.7-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.7-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.6.0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-spark	0.6.0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hudi-presto	0.6.0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.7.1	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.1.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.6.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5.5.64+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.3.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.238.3-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.238.3-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.238.3-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.3	Projet R pour les calculs statistiques
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.6-amzn-0.1	Clients de ligne de commande Spark.
spark-history-server	2.4.6-amzn-0.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.6-amzn-0.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.6-amzn-0.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	2.1.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.31.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.31.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.

Classifications	Description
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.

Classifications	Description
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.

Classifications	Description
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.

Classifications	Description
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.

Classifications	Description
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site .xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.

Classifications	Description
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.31.0

Versions d'application 5.31.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.31.0	emr-5.30.2	emr-5.30.1	emr-5.30.0
AWS SDK pour Java	1,1,852	1,1,759	1,1,759	1,1,759
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,7
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.10.0	1.10.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.6	2.3.6	2.3.6
Hadoop	2.10.0	2.8.5	2.8.5	2.8.5
Hive	2.3.7	2.3.6	2.3.6	2.3.6
Hudi	0.6.0-amzn-0	0.5.2-incubating	0.5.2-incubating	0.5.2-incubating

	emr-5.31.0	emr-5.30.2	emr-5.30.1	emr-5.30.0
Hue	4.7.1	4.6.0	4.6.0	4.6.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.6.0	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,238,3	0.232	0.232	0.232
Spark	2.4.6	2.4.5	2.4.5	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.1.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.31.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.31.0. Les modifications ont été apportées à la version 5.30.1.

Date de parution initiale : 9 octobre 2020

Dernière mise à jour : 15 octobre 2020

Mises à niveau

- Mise à niveau du connecteur Amazon Glue vers la version 1.13.0
- Mise à niveau du SDK Amazon SageMaker Spark vers la version 1.4.0
- Mise à niveau du connecteur Amazon Kinesis vers la version 3.5.9
- Mise à niveau AWS SDK pour Java vers la version 1.11.852
- Mise à niveau de Bigtop-tomcat vers la version 8.5.56
- Mise à niveau d'EMRFS vers la version 2.43.0
- MetricsAndEventsApiGateway Client EMR mis à niveau vers la version 1.4.0
- Mise à niveau d'EMR S3 Dist CP vers la version 2.15.0
- Mise à niveau d'EMR S3 Select vers la version 1.6.0
- Mise à niveau de Flink vers la version 1.11.0
- Mise à niveau de Hadoop vers la version 2.10.0
- Mise à niveau de Hive vers la version 2.3.7
- Mise à niveau de Hudi vers la version 0.6.0
- Mise à niveau de Hue vers la version 4.7.1
- Mise à niveau JupyterHub vers la version 1.1.0
- Mise à niveau de Mxnet vers la version 1.6.0
- Mise à niveau d'OpenCV vers la version 4.3.0
- Mise à niveau de Presto vers la version 0.238.3
- Mise à niveau TensorFlow vers la version 2.1.0

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.

- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Les statistiques des colonnes Hive sont prises en charge pour les versions 5.31.0 et ultérieures d'Amazon EMR.
- Versions de composants mises à niveau.
- Prise en charge d'EMRFS S3EC V2 dans Amazon EMR 5.31.0. Dans les versions 1.11.837 et ultérieures du kit SDK Java S3, le client de chiffrement version 2 (S3EC V2) a été introduit avec diverses améliorations de sécurité. Pour plus d'informations, consultez les ressources suivantes :
  - Article de blog S3 : Updates to the Amazon S3 encryption client.
  - AWS SDK pour Java Guide du développeur : Migrez les clients de chiffrement et de déchiffrement vers la version V2.

- Guide de gestion EMR : Chiffrement côté client Amazon S3.

Le client de chiffrement V1 est toujours disponible dans le kit SDK pour des raisons de rétrocompatibilité.

Nouvelles fonctionnalités

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Redémarrer InstanceController

     ```
     $ sudo systemctl daemon-reload
     ```

```
$ sudo systemctl restart instance-controller
```

Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Avec Amazon EMR 5.31.0, vous pouvez lancer un cluster qui s'intègre à Lake Formation. Cette intégration fournit un filtrage des données précis au niveau des colonnes pour les bases de données et les tables du Glue AWS Data Catalog. Il permet également une authentification unique fédérée à Blocs-notes EMR ou Apache Zeppelin à partir d'un système d'identité d'entreprise. Pour plus d'informations, consultez Intégration d'Amazon EMR avec AWS Lake Formation dans le Guide de gestion Amazon EMR.

  Amazon EMR with Lake Formation est actuellement disponible dans 16 AWS régions : USA Est (Ohio et Virginie du Nord), USA Ouest (Californie du Nord et Oregon), Asie-Pacifique (Mumbai, Séoul, Singapour, Sydney et Tokyo), Canada (Centre), Europe (Francfort, Irlande, Londres, Paris et Stockholm), Amérique du Sud (São Paulo).

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

- Lorsque AtRestEncryption le chiffrement HDFS est activé sur un cluster qui utilise Amazon EMR 5.31.0 ou 5.32.0, les requêtes Hive génèrent l'exception d'exécution suivante.

  ```
  TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
   attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
   java.lang.RuntimeException: Hive Runtime Error while closing
   operators: java.io.IOException: java.util.ServiceConfigurationError:
  ```

```
org.apache.hadoop.security.token.TokenIdentifier: Provider
org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 5.31.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.4.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4,15,0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.13.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.15.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.6.0	Connecteur S3Select EMR
emrfs	2.43.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.11.0	Applications et scripts client de la ligne de commande Apache Flink.
flink-jobmanager-config	1.11.0	Gestion des ressources sur les nœuds EMR pour Apache Flink. JobManager
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.10.0-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.10.0-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.10.0-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.10.0-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.10.0-amzn-0	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.10.0-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.10.0-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.10.0-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.10.0-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.10.0-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.10.0-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.7-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.7-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.7-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.7-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.7-amzn-1	Client Hive-hbase.
hive-metastore-server	2.3.7-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.7-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.6.0-amzn-0	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.

Composant	Version	Description
hudi-spark	0.6.0-amzn-0	Bibliothèque de solution groupée pour exécuter Spark avec Hudi.
hudi-presto	0.6.0-amzn-0	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.7.1	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	1.1.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.6.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5.5.64	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	4.3.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.238.3-amzn-0	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.238.3-amzn-0	Service pour exécuter les éléments d'une requête.
presto-client	0.238.3-amzn-0	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.3	Projet R pour les calculs statistiques

Composant	Version	Description
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.6-amzn-0	Clients de ligne de commande Spark.
spark-history-server	2.4.6-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.6-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.6-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	2.1.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.2	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.31.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.31.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.

Classifications	Description
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.

Classifications	Description
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2

Classifications	Description
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.

Classifications	Description
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.

Classifications	Description
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.

Classifications	Description
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer

Classifications	Description
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.

Classifications	Description
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.30.2

Versions d’application 5.30.2

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.30.2	emr-5.30.1	emr-5.30.0	mr-5.29.0
AWS SDK pour Java	1,1,759	1,1,759	1,1,759	1,11,682
Python	2,7, 3,7	2,7, 3,7	2,7, 3,7	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.10.0	1.10.0	1.9.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.2-incubating	0.5.2-incubating	0.5.0-incubating
Hue	4.6.0	4.6.0	4.6.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.0.0
Livy	0.7.0	0.7.0	0.7.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1

	emr-5.30.2	emr-5.30.1	emr-5.30.0	mr-5.29.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.232	0.232	0.227
Spark	2.4.5	2.4.5	2.4.5	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.30.2

Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.

Modifications, améliorations et problèmes résolus

- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.

- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Le protocole HTTPS est désormais activé par défaut pour les référentiels Amazon Linux. Si vous utilisez une politique VPCE Amazon S3 pour restreindre l'accès à des compartiments spécifiques, vous devez ajouter le nouvel ARN du compartiment Amazon Linux `arn:aws:s3:::amazonlinux-2-repos-$region/*` à votre politique (remplacez `$region` par la région où se trouve le point de terminaison). Pour plus d'informations, consultez cette rubrique dans les forums de AWS discussion. Annonce : Amazon Linux 2 permet désormais d'utiliser le protocole HTTPS lors de la connexion aux référentiels de packages.

Problèmes connus

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :

- Deux partitions ou plus sont analysées à partir de la même table.
- Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
- Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 5.30.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.3.0	Kit de développement logiciel Amazon SageMaker Spark

Composant	Version	Description
emr-ddb	4.14.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.13.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.14.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.5.0	Connecteur S3Select EMR
emrfs	2.40.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.10.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-6.1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-6.1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-6.1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-6.1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.8.5-amzn-6.1	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-6.1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-6.1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-6.1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	2.8.5-amzn-6.1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-6.1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-6.1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.6-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.3.6-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.6-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.6-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.6-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.6-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.6-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.5.2-incubating	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.5.2-incubating	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.6.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
jupyterhub	1.1.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.5.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5.5.64+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.232	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.232	Service pour exécuter les éléments d'une requête.
presto-client	0.232	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.3	Projet R pour les calculs statistiques
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.5-amzn-0.1	Clients de ligne de commande Spark.
spark-history-server	2.4.5-amzn-0.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	2.4.5-amzn-0.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.5-amzn-0.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.14.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.30.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.30.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop

Classifications	Description
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.

Classifications	Description
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.

Classifications	Description
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-note s Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.

Classifications	Description
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password- authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-en v.sh de Presto.

Classifications	Description
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.

Classifications	Description
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.

Classifications	Description
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.30.1

Versions d’application 5.30.1

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.30.1	emr-5.30.0	mr-5.29.0	emr-5.28.1
AWS SDK pour Java	1,1,759	1,1,759	1,11,682	1,1,659
Python	2,7, 3,7	2,7, 3,7	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.10.0	1.9.1	1.9.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.10	1.4.10

	emr-5.30.1	emr-5.30.0	mr-5.29.0	emr-5.28.1
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.2-incubating	0.5.0-incubating	0.5.0-incubating
Hue	4.6.0	4.6.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.232	0.227	0.227
Spark	2.4.5	2.4.5	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.30.1	emr-5.30.0	mr-5.29.0	emr-5.28.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.30.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.30.1. Les modifications ont été apportées à la version 5.30.0.

Date de parution initiale : 30 juin 2020

Dernière mise à jour : 24 août 2020

Modifications, améliorations et problèmes résolus

- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Correction d'un problème à cause duquel le processus du contrôleur d'instance générait un nombre infini de processus.
- Correction d'un problème où Hue ne parvenait pas à exécuter une requête Hive, affichant un message « la base de données est verrouillée » et empêchant l'exécution des requêtes.
- Correction d'un problème lié à Spark qui permettait à un plus grand nombre de tâches de s'exécuter simultanément sur le cluster EMR.
- Correction d'un problème de bloc-notes Jupyter provoquant une erreur « trop de fichiers ouverts » dans le serveur Jupyter.
- Correction d'un problème lié aux temps de démarrage des clusters.

Nouvelles fonctionnalités

- Les interfaces d'application persistantes Tez UI et YARN timeline server sont disponibles avec les versions 6.x d'Amazon EMR et les versions 5.30.1 et ultérieures d'EMR. L'accès par lien en un clic à l'historique des applications persistantes vous permet d'accéder rapidement à l'historique des tâches sans avoir à configurer un proxy web par le biais d'une connexion SSH. Les journaux des clusters actifs et résiliés sont disponibles pendant 30 jours après la fin de l'application. Pour plus d'informations, consultez Affichage des interfaces utilisateur des applications persistantes dans le Guide de gestion Amazon EMR.
- L'exécution de blocs-notes EMR est disponible pour exécuter APIs des blocs-notes EMR via un script ou une ligne de commande. La possibilité de démarrer, d'arrêter, de répertorier et de décrire les exécutions d'un bloc-notes EMR sans la AWS console vous permet de contrôler un bloc-notes EMR par programme. À l'aide d'une cellule de bloc-notes paramétrée, vous pouvez transmettre différentes valeurs de paramètres à un bloc-notes sans avoir à créer une copie du bloc-notes pour chaque nouvel ensemble de valeurs de paramètres. Consultez Actions d'API EMR. Pour un exemple de code, consultez Exemples de commandes pour l'exécution de blocs-notes EMR par programmation.

Problèmes connus

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

Définir explicitement un ulimit à partir de la ligne de commande

1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Redémarrer InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
```

```
sudo systemctl daemon-reload
```

- Blocs-notes EMR

  La fonctionnalité qui permet d'installer des noyaux et des bibliothèques Python supplémentaires sur le nœud primaire du cluster est désactivée par défaut dans la version EMR 5.30.1. Pour plus d'informations sur cette fonctionnalité, consultez Installation de noyaux et de bibliothèques Python sur un nœud primaire de cluster.

  Pour activer cette fonctionnalité, procédez comme suit :

  1. Assurez-vous que la politique d'autorisations attachée à la fonction du service pour les blocs-notes EMR autorise l'action suivante :

     `elasticmapreduce:ListSteps`

     Pour plus d'informations, consultez Rôle de service pour les bloc-notes EMR.

  2. Utilisez le AWS CLI pour exécuter une étape sur le cluster qui configure les Notebooks EMR, comme indiqué dans l'exemple suivant. Remplacez *us-east-1* par la région dans laquelle réside votre cluster. Pour plus d'informations sur l'ajout d'étapes, consultez la rubrique Ajout d'étapes à un cluster à l'aide de la AWS CLI.

     ```
     aws emr add-steps  --cluster-id MyClusterID --steps
      Type=CUSTOM_JAR,Name=EMRNotebooksSetup,ActionOnFailure=CONTINUE,Jar=s3://us-
     east-1.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://
     awssupportdatasvcs.com/bootstrap-actions/EMRNotebooksSetup/emr-notebooks-setup.sh"]
     ```

- Mise à l'échelle gérée

  Les opérations de mise à l'échelle gérées sur des clusters 5.30.0 et 5.30.1 sans Presto installé peuvent provoquer des défaillances d'applications ou empêcher le maintien d'un groupe d'instances ou d'une flotte d'instances uniforme dans l'état `ARRESTED`, en particulier lorsqu'une opération de réduction est rapidement suivie d'une opération d'augementation.

  Pour contourner le problème, choisissez Presto comme application à installer lorsque vous créez un cluster avec les versions 5.30.0 et 5.30.1 d'Amazon EMR, même si votre travail ne nécessite pas Presto.

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

Solution :

- SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
- Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît

entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 5.30.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.3.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.14.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.13.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.14.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.5.0	Connecteur S3Select EMR
emrfs	2.40.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.10.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-6	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».

Composant	Version	Description
hadoop-hdfs-datanode	2.8.5-amzn-6	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-6	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-6	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.8.5-amzn-6	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-6	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-6	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-6	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-6	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-6	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.8.5-amzn-6	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.6-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.6-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.6-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.3.6-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.6-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.6-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.6-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.5.2-incubating	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.5.2-incubating	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.6.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.1.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.5.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5.5.64	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.232	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.232	Service pour exécuter les éléments d'une requête.
presto-client	0.232	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.3	Projet R pour les calculs statistiques
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.5-amzn-0	Clients de ligne de commande Spark.
spark-history-server	2.4.5-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.5-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.5-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	1.14.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.30.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.30.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.

Classifications	Description
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.

Classifications	Description
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.

Classifications	Description
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.

Classifications	Description
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.

Classifications	Description
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site .xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.

Classifications	Description
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.30.0

Versions d’application 5.30.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.30.0	mr-5.29.0	emr-5.28.1	emr-5.28.0
AWS SDK pour Java	1,1,759	1,11,682	1,1,659	1,1,659
Python	2,7, 3,7	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.9.1	1.9.0	1.9.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.0-incubating	0.5.0-incubating	0.5.0-incubating

	emr-5.30.0	mr-5.29.0	emr-5.28.1	emr-5.28.0
Hue	4.6.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.0.0	1.0.0	1.0.0
Livy	0.7.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.227	0.227	0.227
Spark	2.4.5	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.30.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.30.0. Les modifications ont été apportées à la version 5.29.0.

Date de parution initiale : 13 mai 2020

Date de la dernière mise à jour : 25 juin 2020

Mises à niveau

- Mise à niveau AWS SDK pour Java vers la version 1.11.759
- Mise à niveau du SDK Amazon SageMaker Spark vers la version 1.3.0
- Mise à niveau du serveur d'enregistrement EMR vers la version 1.6.0
- Mise à niveau de Flink vers la version 1.10.0
- Mise à niveau de Ganglia vers la version 3.7.2
- Mise à niveau HBase vers la version 1.4.13
- Mise à niveau de Hudi vers la version 0.5.2 incubating
- Mise à niveau de Hue vers la version 4.6.0
- Mise à niveau JupyterHub vers la version 1.1.0
- Mise à niveau de Livy vers la version 0.7.0-incubating
- Mise à niveau d'Oozie vers la version 5.2.0
- Mise à niveau de Presto vers la version 0.232
- Mise à niveau de Spark vers la version 2.4.5
- Connecteurs et pilotes mis à niveau : Amazon Glue Connector 1.12.0 ; Amazon Kinesis Connector 3.5.0 ; EMR DynamoDB Connector 4.14.0

Nouvelles fonctionnalités

- Blocs-notes EMR – Lorsqu'ils sont utilisés avec des clusters EMR créés à l'aide de la version 5.30.0, les noyaux du bloc-notes EMR s'exécutent sur un cluster. Cela améliore les performances des blocs-notes et vous permet d'installer et de personnaliser les noyaux. Vous pouvez également installer des bibliothèques Python sur le nœud primaire du cluster. Pour plus d'informations, consultez Installation et utilisation des noyaux et des bibliothèques dans le Manuel de gestion EMR .

- Dimensionnement géré – Avec les versions 5.30.0 et ultérieures d'Amazon EMR, vous pouvez activer le dimensionnement géré par EMR pour augmenter ou diminuer automatiquement le nombre d'instances ou d'unités dans votre cluster en fonction de la charge de travail. Amazon EMR évalue en permanence les métriques de cluster pour prendre des décisions de dimensionnement qui optimisent vos clusters en termes de coût et de vitesse. Pour plus d'informations, consultez Mise à l'échelle des ressources de cluster dans le Guide de gestion Amazon EMR.
- Chiffrer les fichiers journaux stockés dans Amazon S3 : avec Amazon EMR version 5.30.0 et versions ultérieures, vous pouvez chiffrer les fichiers journaux stockés dans Amazon S3 à l'aide d'une clé gérée par le client. AWS KMS Pour plus d'informations, consultez Chiffrer les fichiers journaux stockés dans Amazon S3 dans le Guide de gestion Amazon EMR.
- Prise en charge d'Amazon Linux 2 – Les versions 5.30.0 et ultérieures d'EMR utilisent le système d'exploitation Amazon Linux 2. La nouvelle personnalisation AMIs (Amazon Machine Image) doit être basée sur l'AMI Amazon Linux 2. Pour en savoir plus, consultez Utilisation d'une image AMI personnalisée.
- Dimensionnement automatique gracieux Presto – Les clusters EMR utilisant la version 5.30.0 peuvent inclure un délai d'attente de dimensionnement automatique qui donne aux tâches Presto le temps s'exécuter complètement avant que leur nœud ne soit hors-service. Pour de plus amples informations, veuillez consulter Utilisation du dimensionnement automatique de Presto avec désaffectation gracieuse.
- Création d'une instance de flotte avec une nouvelle option de stratégie d'allocation – Une nouvelle option de stratégie d'allocation est disponible dans les versions 5.12.1 et ultérieures d'EMR. Cela permet un approvisionnement plus rapide des clusters, une allocation plus précise et moins d'interruptions d'instances Spot. Des mises à jour des rôles de service EMR autres que ceux par défaut sont requises. Consultez Configuration de parcs d'instances.
- Commandes sudo systemctl stop et sudo systemctl start – Dans les versions 5.30.0 et ultérieures d'EMR, qui utilisent le système d'exploitation Amazon Linux 2, EMR utilise les commandes `sudo systemctl stop` et `sudo systemctl start` pour redémarrer les services. Pour plus d'informations, consultez Comment redémarrer un service dans Amazon EMR ?.

Modifications, améliorations et problèmes résolus

- EMR version 5.30.0 n'installe pas Ganglia par défaut. Lorsque vous créez un cluster, vous pouvez sélectionner expressément l’installation de Ganglia.
- Optimisation des performances Spark.
- Optimisation des performances Presto.

- Python 3 est la version par défaut pour Amazon EMR version 5.30.0 et versions ultérieures.
- Le groupe de sécurité géré par défaut pour l'accès au service dans les sous-réseaux privés a été mis à jour avec de nouvelles règles. Si vous utilisez un groupe de sécurité personnalisé pour accéder au service, vous devez inclure les mêmes règles que le groupe de sécurité géré par défaut. Pour plus d'informations, consultez Groupe de sécurité géré par Amazon EMR pour l'accès au service (sous-réseaux privés). Si vous utilisez un rôle de service personnalisé pour Amazon EMR, vous devez accorder l'autorisation aux `ec2:describeSecurityGroups` pour permettre à EMR de confirmer que les groupes de sécurité sont correctement créés. Si vous utilisez le `EMR_DefaultRole`, cette autorisation est déjà incluse dans la stratégie gérée par défaut.

Problèmes connus

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536
     ```

```
LimitNPROC=65536
```

2. Redémarrer InstanceController

```
$ sudo systemctl daemon-reload

$ sudo systemctl restart instance-controller
```

Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Mise à l'échelle gérée

Les opérations de mise à l'échelle gérées sur des clusters 5.30.0 et 5.30.1 sans Presto installé peuvent provoquer des défaillances d'applications ou empêcher le maintien d'un groupe d'instances ou d'une flotte d'instances uniforme dans l'état `ARRESTED`, en particulier lorsqu'une opération de réduction est rapidement suivie d'une opération d'augementation.

Pour contourner le problème, choisissez Presto comme application à installer lorsque vous créez un cluster avec les versions 5.30.0 et 5.30.1 d'Amazon EMR, même si votre travail ne nécessite pas Presto.

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

- Le moteur de base de données par défaut pour Hue 4.6.0 est SQLite, ce qui pose des problèmes lorsque vous essayez d'utiliser Hue avec une base de données externe. Pour résoudre ce problème, définissez `engine` de votre configuration de classification `hue-ini` sur `mysql`. Ce problème a été résolu dans la version 5.30.1 d'Amazon EMR.

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Versions des composants 5.30.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.3.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.14.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.13.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.5.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-notebook-env	1.0.0	Environnement Conda pour bloc-notes emr
emr-s3-dist-cp	2.14.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.5.0	Connecteur S3Select EMR
emrfs	2.40.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.10.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-6	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-6	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-6	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-6	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.8.5-amzn-6	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-6	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-6	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.8.5-amzn-6	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-6	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-6	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-6	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.13	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.13	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.13	HBase client en ligne de commande.
hbase-rest-server	1.4.13	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.4.13	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.6-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.6-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.6-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.6-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.6-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.6-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.6-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.5.2-incubating	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.

Composant	Version	Description
hudi-presto	0.5.2-incubating	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.6.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	1.1.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.7.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.5.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mariadb-server	5.5.64	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.2.0	Client de ligne de commande Oozie.
oozie-server	5.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.

Composant	Version	Description
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.232	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.232	Service pour exécuter les éléments d'une requête.
presto-client	0.232	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.3	Projet R pour les calculs statistiques
ranger-kms-server	1.2.0	Système de gestion des clés Apache Ranger
spark-client	2.4.5-amzn-0	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.4.5-amzn-0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.5-amzn-0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.5-amzn-0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.14.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.30.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.30.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.

Classifications	Description
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.

Classifications	Description
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
hudi-env	Modifiez les valeurs dans l'environnement Hudi.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l'MapReduce application.

Classifications	Description
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.

Classifications	Description
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.

Classifications	Description
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Classifications	Description
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.29.0

Versions d’application 5.29.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	mr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
AWS SDK pour Java	1,11,682	1,1,659	1,1,659	1,1,615
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12

	mr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.1	1.9.0	1.9.0	1.8.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.5
Hudi	0.5.0-incubating	0.5.0-incubating	0.5.0-incubating	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	1.0.0
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	mr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
Presto	0.227	0.227	0.227	0.224
Spark	2.4.4	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.29.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.29.0. Les modifications ont été apportées à la version 5.28.1.

Date de parution initiale : 17 janvier 2020

Mises à niveau

- Mise à niveau AWS SDK pour Java vers la version 1.11.682
- Mise à niveau de Hive vers la version 2.3.6
- Mise à niveau de Flink vers la version 1.9.1
- Mise à niveau d'EmrFS vers la version 2.38.0
- Mise à niveau du connecteur DynamoDB d'EMR vers la version 4.13.0

Modifications, améliorations et problèmes résolus

- Spark
  - Optimisation des performances Spark.

- EMRFS
  - Mise à jour du guide de gestion avec les paramètres par défaut du fichier emrfs-site.xml pour un affichage cohérent.

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.29.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.6	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.13.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.12.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.13.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.4.0	Connecteur S3Select EMR

Composant	Version	Description
emrfs	2.38.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.9.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-5	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-5	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-5	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-5	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	2.8.5-amzn-5	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-5	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-5	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-5	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-5	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-5	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-5	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.10	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.4.10	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.10	HBase client en ligne de commande.
hbase-rest-server	1.4.10	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.10	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.6-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.6-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.6-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.6-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.6-amzn-1	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.6-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.6-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.5.0-incubating	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.5.0-incubating	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.4.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.0.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.6.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.

Composant	Version	Description
mxnet	1.5.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.227	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.227	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
presto-client	0.227	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.4	Clients de ligne de commande Spark.
spark-history-server	2.4.4	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.4	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.4	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.14.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.29.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.29.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.

Classifications	Description
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.

Classifications	Description
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.

Classifications	Description
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-note s Jupyter.

Classifications	Description
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.

Classifications	Description
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.

Classifications	Description
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.

Classifications	Description
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.

Classifications	Description
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.28.1

Versions d’application 5.28.1

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.28.1	emr-5.28.0	emr-5.27.1	emr-5.27.0
AWS SDK pour Java	1,1,659	1,1,659	1,1,615	1,1,615
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.0	1.9.0	1.8.1	1.8.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.5	2.3.5
Hudi	0.5.0-incubating	0.5.0-incubating	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	1.0.0
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.4.0	1.4.0

	emr-5.28.1	emr-5.28.0	emr-5.27.1	emr-5.27.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.2	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.227	0.224	0.224
Spark	2.4.4	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.28.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.28.1. Les modifications ont été apportées à la version 5.28.0.

Date de parution initiale : 10 janvier 2020

Modifications, améliorations et problèmes résolus

- Spark
  - Correction des problèmes de compatibilité Spark.
- CloudWatch Métriques

  - Correction de la publication d'Amazon CloudWatch Metrics sur un cluster EMR comportant plusieurs nœuds principaux.
- Message de journal désactivé
  - Message de faux journal désactivé, « ...en utilisant l'ancienne version (<4.5.8) du client HTTP Apache ».

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.28.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.6	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.12.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.11.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.13.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.3.0	Connecteur S3Select EMR

Composant	Version	Description
emrfs	2.37.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.9.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-5	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-5	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-5	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-5	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	2.8.5-amzn-5	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-5	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-5	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-5	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-5	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-5	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-5	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.10	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.4.10	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.10	HBase client en ligne de commande.
hbase-rest-server	1.4.10	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.10	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.6-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.6-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.6-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.6-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.6-amzn-0	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.6-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.6-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.5.0-incubating	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.5.0-incubating	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.4.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.0.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.6.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.

Composant	Version	Description
mxnet	1.5.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.227	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.227	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
presto-client	0.227	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.4	Clients de ligne de commande Spark.
spark-history-server	2.4.4	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.4	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.4	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.14.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.28.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.28.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.

Classifications	Description
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.

Classifications	Description
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.

Classifications	Description
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-note s Jupyter.

Classifications	Description
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.

Classifications	Description
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.

Classifications	Description
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.

Classifications	Description
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.

Classifications	Description
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.28.0

Versions d’application 5.28.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hudi, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, PigPresto, Spark, Sqoop, TensorFlow, TezZeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.28.0	emr-5.27.1	emr-5.27.0	emr-5.26.0
AWS SDK pour Java	1,1,659	1,1,615	1,1,615	1,1595
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.0	1.8.1	1.8.1	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.5	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.5	2.3.5	2.3.5
Hudi	0.5.0-incubating	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.4.0	1.4.0	1.4.0

	emr-5.28.0	emr-5.27.1	emr-5.27.0	emr-5.26.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.2	4.14.2	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.224	0.224	0.220
Spark	2.4.4	2.4.4	2.4.4	2.4.3
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.13.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.28.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.28.0. Les modifications ont été apportées à la version 5.27.0.

Date de parution initiale : 12 novembre 2019

Mises à niveau

- Mise à niveau de Flink vers la version 1.9.0
- Mise à niveau de Hive vers la version 2.3.6
- Mise à niveau MXNet vers la version 1.5.1
- Mise à niveau de Phoenix vers la version 4.14.3

- Mise à niveau de Presto vers la version 0.227
- Mise à niveau de Zeppelin vers la version 0.8.2

Nouvelles fonctionnalités

- Apache Hudi est désormais disponible pour une installation via Amazon EMR lorsque vous créez un cluster. Pour de plus amples informations, veuillez consulter Hudi.
- (25 novembre 2019) Vous pouvez maintenant choisir d'exécuter plusieurs étapes en parallèle pour améliorer l'utilisation du cluster et faire des économies. Vous pouvez également annuler à la fois les étapes en attente et celles en cours d'exécution. Pour plus d'informations, voir Travailler avec des étapes à l'aide de la console AWS CLI AND.
- (3 décembre 2019) Vous pouvez désormais créer et exécuter des clusters EMR sur. AWS Outposts AWS Outposts permet AWS des services, une infrastructure et des modèles d'exploitation natifs dans les installations sur site. Dans AWS Outposts les environnements, vous pouvez utiliser les mêmes AWS APIs outils et infrastructures que ceux que vous utilisez dans le AWS cloud. Pour plus d'informations, consultez la section Clusters EMR sur. AWS Outposts
- (11 mars 2020) À partir de la version 5.28.0 d'Amazon EMR, vous pouvez créer et exécuter des clusters Amazon EMR sur un sous-réseau de zones AWS locales en tant qu'extension logique d'une région prenant en charge les zones locales. AWS Une zone locale permet aux fonctionnalités d'Amazon EMR et à un sous-ensemble de AWS services, tels que les services de calcul et de stockage, d'être situés plus près des utilisateurs, offrant ainsi un accès à très faible latence aux applications exécutées localement. Pour obtenir la liste des zones locales disponibles, consultez AWS Local Zones. Pour plus d'informations sur l'accès aux zones AWS locales disponibles, voir Régions, zones de disponibilité et zones locales.

  Les zones Locales ne sont actuellement pas compatibles avec Blocs-notes Amazon EMR et ne prennent pas en charge les connexions directes à Amazon EMR via le point de terminaison d'un VPC d'interface (AWS PrivateLink).

Modifications, améliorations et problèmes résolus

- Prise en charge d'applications étendues pour des clusters à haute disponibilité
  - Pour plus d'informations, consultez Applications prises en charge dans un cluster EMR avec plusieurs nœuds primaires dans le Guide de gestion Amazon EMR.
- Spark

  - Optimisation des performances
- Hive
  - Optimisation des performances
- Presto
  - Optimisation des performances

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.28.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.6	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.12.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.11.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.13.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.3.0	Connecteur S3Select EMR

Composant	Version	Description
emrfs	2.37.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.9.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-5	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-5	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-5	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-5	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	2.8.5-amzn-5	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-5	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-5	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-5	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-5	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-5	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-5	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.10	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.4.10	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.10	HBase client en ligne de commande.
hbase-rest-server	1.4.10	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.10	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.6-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.6-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.6-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.6-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.6-amzn-0	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.6-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.6-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hudi	0.5.0-incubating	Infrastructure de traitement incrémentiel pour implémenter un pipeline à faible latence et à efficacité élevée.
hudi-presto	0.5.0-incubating	Bibliothèque de solution groupée pour exécuter Presto avec Hudi.
hue-server	4.4.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.0.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.6.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.

Composant	Version	Description
mxnet	1.5.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,3- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,3- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.227	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.227	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
presto-client	0.227	Client de ligne de commande Presto installé sur les maîtres de secours d'un cluster haute disponibilité où le serveur Presto n'est pas démarré.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.4	Clients de ligne de commande Spark.
spark-history-server	2.4.4	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.4	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.4	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.14.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.28.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.28.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.

Classifications	Description
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.

Classifications	Description
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.

Classifications	Description
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-note s Jupyter.

Classifications	Description
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.

Classifications	Description
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.

Classifications	Description
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.

Classifications	Description
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.

Classifications	Description
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.27.1

Versions d’application 5.27.1

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.27.1	emr-5.27.0	emr-5.26.0	emr-5.25.0
AWS SDK pour Java	1,1,615	1,1,615	1,1595	1,1,566
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.1	1.8.1	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.9
HCatalog	2.3.5	2.3.5	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.5	2.3.5
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0

	emr-5.27.1	emr-5.27.0	emr-5.26.0	emr-5.25.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.2	4.14.2	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.224	0.224	0.220	0.220
Spark	2.4.4	2.4.4	2.4.3	2.4.3
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.13.1	1.13.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notes de mise à jour 5.27.1

Il s'agit d'une version de correctif. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

État de prise en charge du service de métadonnées d'instance (IMDS) V2 : les composants Amazon EMR 5.23.1, 5.27.1 et 5.32 ou versions ultérieures sont utilisés pour tous les appels IMDS. IMDSv2 Pour les appels IMDS dans le code de votre application, vous pouvez utiliser les deux IMDSv1 ou configurer l'IMDS pour qu'il ne soit utilisé que IMDSv2 pour renforcer la sécurité. IMDSv2 Pour les autres versions 5.x EMR, la IMDSv1 désactivation entraîne l'échec du démarrage du cluster.

Versions des composants 5.27.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.4	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.12.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.11.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.13.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.3.0	Connecteur S3Select EMR

Composant	Version	Description
emrfs	2.36.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.8.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	2.8.5-amzn-4	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.10	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.4.10	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.10	HBase client en ligne de commande.
hbase-rest-server	1.4.10	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.10	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.5-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.5-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.5-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.5-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.5-amzn-1	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.5-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.5-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.4.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.0.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.6.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.4.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda

Composant	Version	Description
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,2- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,2- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.224	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.224	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.4	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.4.4	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.4	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.4	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.14.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.27.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.27.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.

Classifications	Description
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.

Classifications	Description
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l'MapReduce application.

Classifications	Description
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.

Classifications	Description
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.

Classifications	Description
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Classifications	Description
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.27.0

Versions d’application 5.27.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
AWS SDK pour Java	1,1,615	1,1595	1,1,566	1,1,546
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.1	1.8.0	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.9	1.4.9
HCatalog	2.3.5	2.3.5	2.3.5	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.5	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	0.9.6	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.2	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
Presto	0.224	0.220	0.220	0.219
Spark	2.4.4	2.4.3	2.4.3	2.4.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.13.1	1.13.1	1.12.0
Tez	0.9.2	0.9.2	0.9.2	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.13

Notes de mise à jour 5.27.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.27.0. Les modifications ont été apportées à la version 5.26.0.

Date de parution initiale : 23 septembre 2019

Mises à niveau

- AWS SDK pour Java 1,1,615
- Flink 1.8.1
- JupyterHub 1.0.0
- Spark 2.4.4
- TensorFlow 1.14.0
- Connecteurs et pilotes :
  - Connecteur DynamoDB 4.12.0

Nouvelles fonctionnalités

- (24 octobre 2019) Les nouvelles fonctionnalités suivantes des blocs-notes EMR sont disponibles avec toutes les versions d'Amazon EMR.
  - Vous pouvez désormais associer des référentiels Git aux blocs-notes EMR pour stocker vos blocs-notes dans un environnement à version contrôlée. Vous pouvez partager du code avec vos pairs et réutiliser les blocs-notes Jupyter existants via des référentiels Git distants. Pour plus d'informations, consultez Associer des référentiels Git à Blocs-notes EMR dans le Guide de gestion Amazon EMR.
  - L'utilitaire nbdime est désormais disponible dans les blocs-notes EMR afin de simplifier la comparaison et la fusion des blocs-notes.
  - Les ordinateurs portables EMR sont désormais compatibles. JupyterLab JupyterLab est un environnement de développement interactif basé sur le Web entièrement compatible avec les ordinateurs portables Jupyter. Vous pouvez désormais choisir d'ouvrir votre bloc-notes dans l'éditeur de bloc-notes Jupyter JupyterLab ou dans l'éditeur de bloc-notes Jupyter.
- (30 octobre 2019) Avec Amazon EMR versions 5.25.0 et ultérieures, vous pouvez vous connecter à l'interface utilisateur du serveur d'historique Spark à partir de la page Récapitulatif du cluster ou de l'onglet Historique de l'application dans la console. Au lieu de configurer un proxy Web par le biais d'une connexion SSH, vous pouvez accéder rapidement à l'interface utilisateur du serveur d'historique Spark pour consulter les métriques de l'application et accéder aux fichiers journaux pertinents pour les clusters actifs et résiliés. Pour plus d'informations, consultez Accès hors cluster aux interfaces utilisateur d'application persistante dans le Guide de gestion Amazon EMR.

Modifications, améliorations et problèmes résolus

- Cluster Amazon EMR avec plusieurs nœuds primaires
  - Vous pouvez installer et exécuter Flink sur un cluster Amazon EMR comportant plusieurs nœuds primaires. Pour plus d'informations, consultez Applications et fonctionnalités prises en charge.
  - Vous pouvez configurer le chiffrement transparent HDFS sur un cluster Amazon EMR comportant plusieurs nœuds primaires. Pour plus d'informations, consultez Chiffrement transparent HDFS sur les clusters EMR comportant plusieurs nœuds primaires.
  - Vous pouvez désormais modifier la configuration des applications exécutées sur un cluster Amazon EMR comportant plusieurs nœuds primaires. Pour plus d'informations, consultez Fourniture d'une configuration pour un groupe d'instances dans un cluster en cours d'exécution.
- Connecteur Amazon EMR-DynamoDB

- Le connecteur Amazon EMR-DynamoDB prend désormais en charge les types de données DynamoDB suivants : boolean, list, map, item, null. Pour plus d'informations, consultez Configuration d'une table Hive pour l'exécution de commandes Hive.

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.27.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.4	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.12.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.11.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.13.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.3.0	Connecteur S3Select EMR

Composant	Version	Description
emrfs	2.36.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.8.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-hdfs-journalnode	2.8.5-amzn-4	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.10	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.4.10	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.10	HBase client en ligne de commande.
hbase-rest-server	1.4.10	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.10	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.5-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.5-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.5-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.5-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.5-amzn-1	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.5-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.5-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.4.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	1.0.0	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.6.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.4.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda

Composant	Version	Description
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,2- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,2- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.224	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.224	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.4	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.4.4	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.4	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.4	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.14.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.27.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.27.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.

Classifications	Description
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.

Classifications	Description
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l'MapReduce application.

Classifications	Description
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.

Classifications	Description
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site.xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.

Classifications	Description
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Classifications	Description
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.26.0

Versions d’application 5.26.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
AWS SDK pour Java	1,1595	1,1,566	1,1,546	1,1,546
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12

	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.9	1.4.9	1.4.9
HCatalog	2.3.5	2.3.5	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
Presto	0.220	0.220	0.219	0.219
Spark	2.4.3	2.4.3	2.4.2	2.4.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.13.1	1.13.1	1.12.0	1.12.0
Tez	0.9.2	0.9.2	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.13	3.4.13

Notes de mise à jour 5.26.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.26.0. Les modifications ont été apportées à la version 5.25.0.

Date de parution initiale : 8 août 2019

Dernière mise à jour : 19 août 2019

Mises à niveau

- AWS SDK pour Java 1,1595
- HBase 1.4.10
- Phoenix 4.14.2
- Connecteurs et pilotes :
  - Connecteur DynamoDB 4.11.0
  - Connecteur MariaDB 2.4.2
  - Pilote JDBC Amazon Redshift version 1.2.32.1056

Nouvelles fonctionnalités

- (Bêta) Avec Amazon EMR 5.26.0, vous pouvez lancer un cluster qui s'intègre à Lake Formation. Cette intégration fournit un accès détaillé au niveau des colonnes aux bases de données et aux tables du Glue AWS Data Catalog. Il permet également une authentification unique fédérée à Blocs-notes EMR ou Apache Zeppelin à partir d'un système d'identité d'entreprise. Pour plus d'informations, consultez Intégrer Amazon EMR à AWS Lake Formation (version bêta).
- (19 août 2019) Bloquer l'accès public d'Amazon EMR est désormais disponible avec toutes les versions d'Amazon EMR qui prennent en charge les groupes de sécurité. Bloquer l'accès public est un paramètre applicable à l'ensemble du compte et appliqué à chaque AWS région. Le blocage de l'accès public empêche le lancement d'un cluster lorsqu'un groupe de sécurité associé au cluster possède une règle autorisant le trafic entrant depuis IPv4 0.0.0.0/0 ou IPv6 : :/0 (accès public) sur un port, sauf si un port est spécifié comme exception. Le port 22 est une exception par défaut. Pour plus d'informations, consultez Utilisation d'Amazon EMR Block Public Access dans le Guide de gestion Amazon EMR.

Modifications, améliorations et problèmes résolus

- Blocs-notes EMR
  - Avec EMR 5.26.0 et versions ultérieures, Blocs-notes EMR prend en charge les bibliothèques Python limitées aux blocs-notes en plus des bibliothèques Python par défaut. Vous pouvez installer des bibliothèques limitées au bloc-notes depuis l'éditeur de bloc-notes sans avoir à recréer un cluster ou à rattacher un bloc-notes à un cluster. Les bibliothèques limitées aux blocs-notes sont créées dans un environnement virtuel Python, de sorte qu'elles ne s'appliquent qu'à la session de bloc-notes en cours. Cela vous permet d'isoler les dépendances du bloc-notes. Pour plus d'informations, consultez Utilisation de bibliothèques limitées aux blocs-notes dans le Guide de gestion Amazon EMR.
- EMRFS
  - Vous pouvez activer une fonctionnalité ETag de vérification (version bêta) en réglant `fs.s3.consistent.metadata.etag.verification.enabled` sur`true`. Grâce à cette fonctionnalité, EMRFS utilise Amazon S3 ETags pour vérifier que les objets lus sont de la dernière version disponible. Cette fonctionnalité est utile dans les cas read-after-update d'utilisation dans lesquels des fichiers sur Amazon S3 sont remplacés tout en conservant le même nom. Cette fonctionnalité de ETag vérification ne fonctionne pas actuellement avec S3 Select. Pour plus d'informations, consultez Configuration de la vue cohérente.
- Spark

  - Les optimisations suivantes sont désormais activées par défaut : élimination dynamique des partitions, DISTINCT avant INTERSECT, amélioration de l'inférence des statistiques de plan SQL pour les requêtes JOIN suivies de DISTINCT, aplatissement des sous-requêtes scalaires, réorganisation optimisée des jointures et jointure par filtre bloom. Pour plus d'informations, consultez Optimisation des performances de Spark.
  - Génération de code améliorée pour l'ensemble de l'étape pour Sort Merge Join.
  - Amélioration de la réutilisation des fragments de requête et des sous-requêtes.
  - Améliorations apportées à la préallocation des exécuteurs au démarrage de Spark.
  - Les jointures par filtre bloom ne sont plus appliquées lorsque le côté le plus petit de la jointure inclut un indice de diffusion.
- Tez
  - Un problème avec Tez a été résolu. L'interface utilisateur de Tez fonctionne désormais sur un cluster Amazon EMR avec plusieurs nœuds primaires.

Problèmes connus

- Les fonctionnalités améliorées de génération de code pour l'ensemble de l'étape pour Sort Merge Join peuvent augmenter la sollicitation de la mémoire lorsqu'elles sont activées. Cette optimisation améliore les performances, mais peut entraîner de nouvelles tentatives de tâches ou des échecs si `spark.yarn.executor.memoryOverheadFactor` n'est pas réglé pour fournir suffisamment de mémoire. Pour désactiver cette fonctionnalité, définissez `spark.sql.sortMergeJoinExec.extendedCodegen.enabled` sur false.
- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

- SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
- Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.26.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.4	Kit de développement logiciel Amazon SageMaker Spark

Composant	Version	Description
emr-ddb	4.11.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.10.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.12.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.3.0	Connecteur S3Select EMR
emrfs	2.35.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.8.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.8.5-amzn-4	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	2.8.5-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.10	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.10	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.10	HBase client en ligne de commande.
hbase-rest-server	1.4.10	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.10	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.5-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.3.5-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.5-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.5-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.5-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.5-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.5-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.4.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	0.9.6	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.6.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.4.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14,2- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14,2- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.220	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.220	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.3	Clients de ligne de commande Spark.
spark-history-server	2.4.3	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.3	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.3	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.13.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.8.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.26.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.26.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.

Classifications	Description
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.

Classifications	Description
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.

Classifications	Description
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.

Classifications	Description
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.

Classifications	Description
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.25.0

Versions d’application 5.25.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
AWS SDK pour Java	1,1,566	1,1,546	1,1,546	1,1,519
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.8.0	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
HCatalog	2.3.5	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.6	0.9.4
Livy	0.6.0	0.6.0	0.6.0	0,5,0
MXNet	1.4.0	1.4.0	1.4.0	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.220	0.219	0.219	0.215
Spark	2.4.3	2.4.2	2.4.2	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.13.1	1.12.0	1.12.0	1.12.0
Tez	0.9.2	0.9.1	0.9.1	0.9.1

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.13	3.4.13	3.4.13

Notes de mise à jour 5.25.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.25.0. Les modifications ont été apportées à la version 5.24.1.

Date de parution initiale : 17 juillet 2019

Dernière mise à jour : 30 octobre 2019

Amazon EMR 5.25.0

Mises à niveau

- AWS SDK pour Java 1,1,566
- Hive 2.3.5
- Presto 0.220
- Spark 2.4.3
- TensorFlow 1.13.1
- Tez 0.9.2
- Zookeeper 3.4.14

Nouvelles fonctionnalités

- (30 octobre 2019) À partir de la version 5.25.0 d'Amazon EMR, vous pouvez vous connecter à l'interface utilisateur du serveur d'historique Spark à partir de la page Récapitulatif du cluster ou de l'onglet Historique de l'application dans la console. Au lieu de configurer un proxy Web par le biais d'une connexion SSH, vous pouvez accéder rapidement à l'interface utilisateur du serveur d'historique Spark pour consulter les métriques de l'application et accéder aux fichiers journaux

pertinents pour les clusters actifs et résiliés. Pour plus d'informations, consultez Accès hors cluster aux interfaces utilisateur d'application persistante dans le Guide de gestion Amazon EMR.

Modifications, améliorations et problèmes résolus

- Spark
  - Amélioration des performances de certaines jointures en utilisant des filtres Bloom pour préfiltrer les entrées. L'optimisation est désactivée par défaut et peut être activée en définissant le paramètre de configuration Spark `spark.sql.bloomFilterJoin.enabled` sur `true`.
  - Amélioration des performances du regroupement par colonnes de type chaîne.
  - Amélioration de la mémoire par défaut de l'exécuteur Spark et de la configuration des cœurs des types d'instances R4 pour les clusters non HBase installés.
  - Résolution d'un problème antérieur lié à la fonctionnalité d'élimination dynamique des partitions, où la table éliminée devait se trouver du côté gauche de la jointure.
  - Amélioration de l'optimisation DISTINCT avant INTERSECT pour l'appliquer à des cas supplémentaires impliquant des alias.
  - Amélioration de l'inférence des statistiques de plan SQL pour les requêtes JOIN suivies de DISTINCT. Cette amélioration est désactivée par défaut et peut être activée en définissant le paramètre de configuration Spark `spark.sql.statsImprovements.enabled` sur `true`. Cette optimisation est requise par la fonctionnalité Distinct avant Intersect et sera automatiquement activée lorsque `spark.sql.optimizer.distinctBeforeIntersect.enabled` est définie sur `true`.
  - Ordre de jointure optimisé en fonction de la taille de la table et des filtres. Cette optimisation est désactivée par défaut et peut être activée en définissant le paramètre de configuration Spark `spark.sql.optimizer.sizeBasedJoinReorder.enabled` sur `true`.

  Pour plus d'informations, consultez Optimisation des performances de Spark.
- EMRFS
  - Le paramètre EMRFS, `fs.s3.buckets.create.enabled`, est désormais désactivé par défaut. Lors des tests, nous avons constaté que la désactivation de ce paramètre améliore les performances et empêche la création involontaire de compartiments S3. Si votre application repose sur cette fonctionnalité, vous pouvez l'activer en définissant la propriété `fs.s3.buckets.create.enabled` sur `true` dans la classification de configuration `emrfs-site`. Pour plus d'informations, consultez Fourniture d'une configuration lors de la création d'un cluster.

- Améliorations du chiffrement du disque local et du chiffrement S3 dans les configurations de sécurité (5 août 2019)
  - Les paramètres de chiffrement Amazon S3 ont été séparés des paramètres de chiffrement du disque local dans la configuration de la sécurité.
  - Ajout d'une option permettant d'activer le chiffrement EBS avec les versions 5.24.0 et ultérieures. La sélection de cette option chiffre le volume du périphérique racine en plus des volumes de stockage. Les versions précédentes nécessitaient l'utilisation d'une AMI personnalisée pour chiffrer le volume du périphérique racine.
  - Pour plus d'informations, consultez Options de chiffrement dans le Guide de gestion Amazon EMR.

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.25.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.4	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.10.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.9.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.11.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.3.0	Connecteur S3Select EMR
emrfs	2.34.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.8.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».

Composant	Version	Description
hadoop-hdfs-datanode	2.8.5-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.8.5-amzn-4	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.8.5-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.9	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.9	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.9	HBase client en ligne de commande.
hbase-rest-server	1.4.9	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.9	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.5-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.5-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.5-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.3.5-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.5-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.5-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.5-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.4.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	0.9.6	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.6.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.4.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.

Composant	Version	Description
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.1- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.1- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.220	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.220	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques

Composant	Version	Description
spark-client	2.4.3	Clients de ligne de commande Spark.
spark-history-server	2.4.3	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.3	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.3	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.13.1	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.2	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.1	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.14	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.14	ZooKeeper client en ligne de commande.

Classifications des configurations 5.25.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.25.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.

Classifications	Description
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.

Classifications	Description
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2

Classifications	Description
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.

Classifications	Description
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.

Classifications	Description
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
recordserver-env	Modifiez les valeurs dans l'environnement EMR. RecordServer
recordserver-conf	Modifiez les valeurs dans le fichier erver.properties RecordServer de l'EMR.
recordserver-log4j	Modifiez les valeurs dans le fichier log4j.properties RecordServer de l'EMR.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark

Classifications	Description
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.24.1

Versions d’application 5.24.1

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.24.1	emr-5.24.0	emr-5.23.1	emr-5.23.0
AWS SDK pour Java	1,1,546	1,1,546	1,1,519	1,1,519
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.7.1	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9

	emr-5.24.1	emr-5.24.0	emr-5.23.1	emr-5.23.0
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.4	0.9.4
Livy	0.6.0	0.6.0	0,5,0	0,5,0
MXNet	1.4.0	1.4.0	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.219	0.219	0.215	0.215
Spark	2.4.2	2.4.2	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1

	emr-5.24.1	emr-5.24.0	emr-5.23.1	emr-5.23.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notes de mise à jour 5.24.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.24.1. Les modifications ont été apportées à la version 5.24.0.

Date de parution initiale : 26 juin 2019

Modifications, améliorations et problèmes résolus

- Mise à jour de l'AMI Amazon Linux par défaut pour Amazon EMR afin d'inclure d'importantes mises à jour de sécurité du noyau Linux, notamment le problème de déni de service TCP SACK (AWS-2019-005).

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.

- Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

  ```
  kinit -kt <keytab_file> <principal>
  ```

  Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.24.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.1	Kit de développement logiciel Amazon SageMaker Spark

Composant	Version	Description
emr-ddb	4.9.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.8.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.11.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.3.0	Connecteur S3Select EMR
emrfs	2.33.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.8.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.8.5-amzn-4	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	2.8.5-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.9	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.9	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.9	HBase client en ligne de commande.
hbase-rest-server	1.4.9	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.9	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.4-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.3.4-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.4-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.4-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.4-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.4-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.4-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.4.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	0.9.6	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.6.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.4.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.1- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.1- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.219	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.219	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.2	Clients de ligne de commande Spark.
spark-history-server	2.4.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.12.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.1	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.8.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.13	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.13	ZooKeeper client en ligne de commande.

Classifications des configurations 5.24.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.24.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.

Classifications	Description
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.

Classifications	Description
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.

Classifications	Description
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.

Classifications	Description
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.

Classifications	Description
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.24.0

Versions d'application 5.24.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.24.0	emr-5.23.1	emr-5.23.0	emr-5.22.0
AWS SDK pour Java	1,1,546	1,1,519	1,1,519	1,11,510
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.7.1	1.7.1	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-

	emr-5.24.0	emr-5.23.1	emr-5.23.0	emr-5.22.0
Hue	4.4.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.4	0.9.4	0.9.4
Livy	0.6.0	0,5,0	0,5,0	0,5,0
MXNet	1.4.0	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.219	0.215	0.215	0.215
Spark	2.4.2	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notes de mise à jour 5.24.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.24.0. Les modifications ont été apportées à la version 5.23.0.

Date de parution initiale : 11 juin 2019

Dernière mise à jour : 5 août 2019

Mises à niveau

- Flink 1.8.0
- Hue 4.4.0
- JupyterHub 0,9,6
- Livy 0.6.0
- MxNet 1.4.0
- Presto 0.219
- Spark 2.4.2
- AWS SDK pour Java 1,1,546
- Connecteurs et pilotes :
  - Connecteur DynamoDB 4.9.0
  - Connecteur MariaDB 2.4.1
  - Pilote JDBC Amazon Redshift version 1.2.27.1051

Modifications, améliorations et problèmes résolus

- Spark
  - Ajout d'une optimisation pour éliminer dynamiquement les partitions. L'optimisation est désactivée par défaut. Pour l'activer, définissez le paramètre de configuration Spark `spark.sql.dynamicPartitionPruning.enabled` sur `true`.
  - Amélioration des performances des requêtes `INTERSECT`. L'optimisation est désactivée par défaut. Pour l'activer, définissez le paramètre de configuration Spark `spark.sql.optimizer.distinctBeforeIntersect.enabled` sur `true`.
  - Ajout d'une optimisation pour aplatir les sous-requêtes scalaires avec des agrégats qui utilisent la même relation. L'optimisation est désactivée

par défaut. Pour l'activer, définissez le paramètre de configuration Spark `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled` sur `true`.

- Génération de code améliorée pour l'ensemble de l'étape.

Pour plus d'informations, consultez Optimisation des performances de Spark.

- Améliorations du chiffrement du disque local et du chiffrement S3 dans les configurations de sécurité (5 août 2019)
  - Les paramètres de chiffrement Amazon S3 ont été séparés des paramètres de chiffrement du disque local dans la configuration de la sécurité.
  - Ajout d'une option pour activer le chiffrement EBS. La sélection de cette option chiffre le volume du périphérique racine en plus des volumes de stockage. Les versions précédentes nécessitaient l'utilisation d'une AMI personnalisée pour chiffrer le volume du périphérique racine.
  - Pour plus d'informations, consultez Options de chiffrement dans le Guide de gestion Amazon EMR.

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.24.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.9.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-goodies	2.8.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.11.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.3.0	Connecteur S3Select EMR
emrfs	2.33.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.8.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.8.5-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.8.5-amzn-4	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.8.5-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.9	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.9	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.9	HBase client en ligne de commande.
hbase-rest-server	1.4.9	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.9	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.4-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.3.4-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.4-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.4-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.4-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.4-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.4-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.4.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	0.9.6	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.6.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.4.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.1- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.1- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.219	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.219	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.2	Clients de ligne de commande Spark.
spark-history-server	2.4.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.12.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.1	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.8.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.13	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.13	ZooKeeper client en ligne de commande.

Classifications des configurations 5.24.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.24.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.

Classifications	Description
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.

Classifications	Description
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.

Classifications	Description
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.

Classifications	Description
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.

Classifications	Description
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.23.1

Versions d’application 5.23.1

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.23.1	emr-5.23.0	emr-5.22.0	emr-5.21.2
AWS SDK pour Java	1,1,519	1,1,519	1,11,510	1,11,479
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.1	1.7.1	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-

	emr-5.23.1	emr-5.23.0	emr-5.22.0	emr-5.21.2
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.0.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notes de mise à jour 5.23.1

Il s'agit d'une version de correctif. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

État de prise en charge du service de métadonnées d'instance (IMDS) V2 : les composants Amazon EMR 5.23.1, 5.27.1 et 5.32 ou versions ultérieures sont utilisés pour tous les appels IMDS. IMDSv2 Pour les appels IMDS dans le code de votre application, vous pouvez utiliser les deux IMDSv1 ou configurer l'IMDS pour qu'il ne soit utilisé que IMDSv2 pour renforcer la sécurité. IMDSv2 Pour les autres versions 5.x EMR, la IMDSv1 désactivation entraîne l'échec du démarrage du cluster.

Versions des composants 5.23.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.8.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.7.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.11.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.2.0	Connecteur S3Select EMR
emrfs	2.32.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.7.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».

Composant	Version	Description
hadoop-hdfs-datanode	2.8.5-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.8.5-amzn-3	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.8.5-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.9	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.9	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.9	HBase client en ligne de commande.
hbase-rest-server	1.4.9	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.9	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.4-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.4-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.4-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.3.4-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.4-amzn-1	Client Hive-hbase.
hive-metastore-server	2.3.4-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.4-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.3.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	0.9.4	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.3.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.

Composant	Version	Description
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.1- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.1- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.215	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.215	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques

Composant	Version	Description
spark-client	2.4.0	Clients de ligne de commande Spark.
spark-history-server	2.4.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.12.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.1	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.1	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.13	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.13	ZooKeeper client en ligne de commande.

Classifications des configurations 5.23.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.23.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.

Classifications	Description
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.

Classifications	Description
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2

Classifications	Description
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.

Classifications	Description
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.

Classifications	Description
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.

Classifications	Description
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.23.0

Versions d’application 5.23.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.23.0	emr-5.22.0	emr-5.21.2	emr-5.21.1
AWS SDK pour Java	1,1,519	1,11,510	1,11,479	1,11,479
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.1	1.7.0	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-

	emr-5.23.0	emr-5.22.0	emr-5.21.2	emr-5.21.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.0.0	5.0.0
Phoenix	4.14.1	4.14.1	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notes de mise à jour 5.23.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.23.0. Les modifications ont été apportées à la version 5.22.0.

Date de parution initiale : 1 avril 2019

Dernière mise à jour : 30 avril 2019

Mises à niveau

- AWS SDK pour Java 1,1,519

Nouvelles fonctionnalités

- (30 avril 2019) Avec Amazon EMR 5.23.0 et versions ultérieures, vous pouvez lancer un cluster avec trois nœuds principaux pour prendre en charge la haute disponibilité d'applications telles que YARN Resource Manager, HDFS, Spark NameNode, Hive et Ganglia. Le nœud primaire n'est plus un point de défaillance potentiel grâce à cette fonctionnalité. Si l'un des nœuds primaires tombe en panne, Amazon EMR passe automatiquement sur un nœud primaire de secours et remplace le nœud primaire défaillant par un nouveau nœud ayant la même configuration et les mêmes actions de démarrage. Pour plus d'informations, consultez Planification et configuration des nœuds primaires.

Problèmes connus

- Interface utilisateur Tez (corrigée dans la version 5.26.0 d'Amazon EMR)

  L'interface utilisateur de Tez ne fonctionne pas sur un cluster EMR comportant plusieurs nœuds primaires.
- Hue (corrigée dans la version 5.24.0 d'Amazon EMR)
  - Hue exécuté sur Amazon EMR ne prend pas en charge Solr. À compter de la version 5.20.0 d'Amazon EMR, un problème de configuration incorrect entraîne l'activation de Solr et un message d'erreur inoffensif semblable au suivant s'affiche :

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Pour empêcher l'affichage du message d'erreur de Solr, procédez comme suit :

1. Connectez-vous à la ligne de commande du nœud primaire à l'aide de SSH.
2. Utilisez un éditeur de texte pour ouvrir le fichier `hue.ini`. Par exemple :

   ```
   sudo vim /etc/hue/conf/hue.ini
   ```

3. Recherchez le terme `appblacklist` et modifiez la ligne comme suit :

   ```
   appblacklist = search
   ```

4. Enregistrez vos modifications et redémarrez Hue comme indiqué dans l'exemple suivant :

   ```
   sudo stop hue; sudo start hue
   ```

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos

paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.23.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.8.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.7.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-s3-dist-cp	2.11.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.2.0	Connecteur S3Select EMR
emrfs	2.32.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.7.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-3	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	2.8.5-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-hdfs-journalnode	2.8.5-amzn-3	Service HDFS pour la gestion du journal du système de fichiers Hadoop sur les clusters HA.
hadoop-httpfs-server	2.8.5-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	1.4.9	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.9	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.9	HBase client en ligne de commande.
hbase-rest-server	1.4.9	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.9	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.4-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.4-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.4-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.4-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.4-amzn-1	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.4-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.4-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.3.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	0.9.4	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.3.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda

Composant	Version	Description
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.1- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.1- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.215	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.215	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.0	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.4.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.12.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.1	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.13	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.13	ZooKeeper client en ligne de commande.

Classifications des configurations 5.23.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.23.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.

Classifications	Description
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.

Classifications	Description
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l'MapReduce application.

Classifications	Description
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.

Classifications	Description
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.

Classifications	Description
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.22.0

Versions d'application 5.22.0

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR

- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.22.0	emr-5.21.2	emr-5.21.1	emr-5.21.0
AWS SDK pour Java	1,11,510	1,11,479	1,11,479	1,11,479
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.0	1.7.0	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.8	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4

	emr-5.22.0	emr-5.21.2	emr-5.21.1	emr-5.21.0
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.1	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notes de mise à jour 5.22.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.22.0. Les modifications ont été apportées à la version 5.21.0.

Important

Depuis la version 5.22.0 d'Amazon EMR, Amazon EMR AWS utilise Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Les versions antérieures d'Amazon EMR utilisent la version 2 de AWS Signature dans certains cas, sauf

si les notes de publication indiquent que la version 4 de Signature est utilisée exclusivement. Pour plus d'informations, consultez les sections Authentification des demandes (AWS Signature version 4) et Authentification des demandes (AWS Signature version 2) dans le manuel Amazon Simple Storage Service Developer Guide.

Date de parution initiale : 20 mars 2019

Mises à niveau

- Flink 1.7.1
- HBase 1.4.9
- Oozie 5.1.0
- Phoenix 4.14.1
- Zeppelin 0.8.1
- Connecteurs et pilotes :
  - Connecteur DynamoDB 4.8.0
  - Connecteur MariaDB 2.2.6
  - Pilote JDBC Amazon Redshift version 1.2.20.1043

Nouvelles fonctionnalités

- Modification de la configuration EBS par défaut pour les types d' EC2 instances avec stockage EBS uniquement. Lorsque vous créez un cluster à l'aide d'Amazon EMR version 5.22.0 et versions ultérieures, la quantité de stockage EBS par défaut augmente en fonction de la taille de l'instance. En outre, nous avons fractionné une plus grande quantité de stockage sur plusieurs volumes, afin d'offrir de meilleures performances en termes d'IOPS. Si vous voulez utiliser une autre configuration de stockage d'instance EBS, vous pouvez la spécifier lorsque vous créez un cluster EMR ou ajoutez des nœuds à un cluster existant. Pour plus d'informations sur la quantité de stockage et le nombre de volumes alloués par défaut pour chaque type d'instance, consultez Stockage EBS par défaut pour les instances dans le Guide de gestion Amazon EMR.

Modifications, améliorations et problèmes résolus

- Spark

  - Introduction d'une nouvelle propriété de configuration pour Spark sur YARN, `spark.yarn.executor.memoryOverheadFactor`. La valeur de cette propriété est un facteur d'échelle qui définit la valeur de la surcharge mémoire à un pourcentage de la mémoire de l'exécuteur, avec un minimum de 384 Mo. Si la surcharge mémoire est définie explicitement à l'aide `spark.yarn.executor.memoryOverhead`, cette propriété n'a aucun effet. La valeur par défaut est `0.1875`, ce qui représente 18,75 %. Cette valeur par défaut pour Amazon EMR laisse plus d'espace dans les conteneurs YARN pour la surcharge de mémoire de l'exécuteur que la valeur par défaut de 10 % définie en interne par Spark. La valeur par défaut d'Amazon EMR de 18,75 % a montré de façon empirique moins de défaillances liées à la mémoire dans les évaluations comparatives TPC-DS.
  - Rétroportage de SPARK-26316 pour améliorer les performances.
- Dans les versions 5.19.0, 5.20.0 et 5.21.0 d'Amazon EMR, les étiquettes des nœuds YARN sont stockées dans un répertoire HDFS. Dans certaines situations, cela entraîne des retards dans le démarrage des nœuds principaux, ce qui provoque le dépassement du délai du cluster et l'échec du lancement. À partir d'Amazon EMR 5.22.0, ce problème est résolu. Les étiquettes des nœuds YARN sont stockées sur le disque local de chaque nœud du cluster, évitant ainsi toute dépendance vis-à-vis du HDFS.

Problèmes connus

- Hue (corrigée dans la version 5.24.0 d'Amazon EMR)
  - Hue exécuté sur Amazon EMR ne prend pas en charge Solr. À compter de la version 5.20.0 d'Amazon EMR, un problème de configuration incorrect entraîne l'activation de Solr et un message d'erreur inoffensif semblable au suivant s'affiche :

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Pour empêcher l'affichage du message d'erreur de Solr, procédez comme suit :

    1. Connectez-vous à la ligne de commande du nœud primaire à l'aide de SSH.
    2. Utilisez un éditeur de texte pour ouvrir le fichier `hue.ini`. Par exemple :

```
sudo vim /etc/hue/conf/hue.ini
```

3. Recherchez le terme `appblacklist` et modifiez la ligne comme suit :

```
appblacklist = search
```

4. Enregistrez vos modifications et redémarrez Hue comme indiqué dans l'exemple suivant :

```
sudo stop hue; sudo start hue
```

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

  Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.22.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.8.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.6.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.11.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.2.0	Connecteur S3Select EMR

Composant	Version	Description
emrfs	2.31.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.7.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-httpfs-server	2.8.5-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.9	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.9	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.9	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	1.4.9	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.9	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.4-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.4-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.4-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.4-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.4-amzn-1	Client Hive-hbase.
hive-metastore-server	2.3.4-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.4-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.

Composant	Version	Description
hue-server	4.3.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	0.9.4	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.3.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.1.0	Client de ligne de commande Oozie.
oozie-server	5.1.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur

Composant	Version	Description
phoenix-library	4,14.1- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.1- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.215	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.215	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.0	Clients de ligne de commande Spark.
spark-history-server	2.4.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.0	Bibliothèques Apache Spark requises par les esclaves YARN.

Composant	Version	Description
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.12.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.1	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.13	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.13	ZooKeeper client en ligne de commande.

Classifications des configurations 5.22.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.22.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.

Classifications	Description
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.

Classifications	Description
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.

Classifications	Description
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.

Classifications	Description
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.

Classifications	Description
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.

Classifications	Description
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.21.2

Versions d’application 5.21.2

Cette version inclut les applications suivantes : FlinkGangliaHBase, HCatalog, HadoopHive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, PrestoSpark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
AWS SDK pour Java	1,11,479	1,11,479	1,11,479	1,11,461
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.7.0	1.7.0	1.6.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.214
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notes de mise à jour 5.21.2

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.21.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent

généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.7.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.5.1	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.11.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.2.0	Connecteur S3Select EMR
emrfs	2.30.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.7.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.5-amzn-1	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.8.5-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.8	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.8	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.8	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	1.4.8	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.8	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.4-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.4-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.4-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.4-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.4-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.4-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.4-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.

Composant	Version	Description
hue-server	4.3.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	0.9.4	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.3.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur

Composant	Version	Description
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.215	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.215	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.0	Clients de ligne de commande Spark.
spark-history-server	2.4.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.0	Bibliothèques Apache Spark requises par les esclaves YARN.

Composant	Version	Description
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.12.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.1	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.13	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.13	ZooKeeper client en ligne de commande.

Classifications des configurations 5.21.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.21.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.

Classifications	Description
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.

Classifications	Description
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.

Classifications	Description
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.

Classifications	Description
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.

Classifications	Description
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.

Classifications	Description
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.21.1

Versions d’application 5.21.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
AWS SDK pour Java	1,11,479	1,11,479	1,11,461	1,11,461
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,6

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
Scala	2,1,12	2,1,12	2,1,12	2,1,12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.7.0	1.6.2	1.6.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.214	0.214
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notes de mise à jour 5.21.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.21.1. Les modifications ont été apportées à la version 5.21.0.

Date de parution initiale : 18 juillet 2019

Modifications, améliorations et problèmes résolus

- Mise à jour de l'AMI Amazon Linux par défaut pour Amazon EMR afin d'inclure d'importantes mises à jour de sécurité du noyau Linux, notamment le problème de déni de service TCP SACK (AWS-2019-005).

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

Solution :

- SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
- Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.21.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.7.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.5.1	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.11.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.2.0	Connecteur S3Select EMR
emrfs	2.30.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.7.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst

Composant	Version	Description
		ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.5-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.8	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.8	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.8	HBase client en ligne de commande.
hbase-rest-server	1.4.8	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.4.8	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.4-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.4-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.4-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.4-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.4-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.4-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.4-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.3.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
jupyterhub	0.9.4	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.3.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.215	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.215	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.0	Clients de ligne de commande Spark.
spark-history-server	2.4.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.

Composant	Version	Description
tensorflow	1.12.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.1	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.13	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.13	ZooKeeper client en ligne de commande.

Classifications des configurations 5.21.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.21.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.

Classifications	Description
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.

Classifications	Description
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.

Classifications	Description
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.

Classifications	Description
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Classifications	Description
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.21.0

Versions d’application 5.21.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.21.0	emr-5.20.1	emr-5.20.0	emr-5.19.1
AWS SDK pour Java	1,11,479	1,11,461	1,11,461	1,1,433
Python	2,7, 3,6	2,7, 3,6	2,7, 3,6	2,7, 3,4
Scala	2,1,12	2,1,12	2,1,12	2.11.8

	emr-5.21.0	emr-5.20.1	emr-5.20.0	emr-5.19.1
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.6.2	1.6.2	1.6.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.7
HCatalog	2.3.4	2.3.4	2.3.4	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.3
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.3.1	1.3.1	1.3.1	1.3.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-5.21.0	emr-5.20.1	emr-5.20.0	emr-5.19.1
Presto	0.215	0.214	0.214	0.212
Spark	2.4.0	2.4.0	2.4.0	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.11.0
Tez	0.9.1	0.9.1	0.9.1	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notes de mise à jour 5.21.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.21.0. Les modifications ont été apportées à la version 5.20.0.

Date de parution initiale : 18 février 2019

Dernière mise à jour : 3 avril 2019

Mises à niveau

- Flink 1.7.0
- Presto 0.215
- AWS SDK pour Java 1,11,479

Nouvelles fonctionnalités

- (3 avril 2019) Avec la version 5.21.0 et ultérieures d'Amazon EMR, vous permet de remplacer les configurations de cluster et de spécifier des classifications de configuration supplémentaires pour chaque groupe d'instances dans un cluster en cours d'exécution. Pour ce faire, utilisez la console Amazon EMR, le AWS Command Line Interface (AWS CLI) ou le AWS SDK. Pour plus

d'informations, consultez Fourniture d'une configuration pour un groupe d'instances dans un cluster en cours d'exécution.

Modifications, améliorations et problèmes résolus

- Zeppelin
  - Rétroportage de ZEPPELIN-3878.

Problèmes connus

- Hue (corrigée dans la version 5.24.0 d'Amazon EMR)
  - Hue exécuté sur Amazon EMR ne prend pas en charge Solr. À compter de la version 5.20.0 d'Amazon EMR, un problème de configuration incorrect entraîne l'activation de Solr et un message d'erreur inoffensif semblable au suivant s'affiche :

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Pour empêcher l'affichage du message d'erreur de Solr, procédez comme suit :

    1. Connectez-vous à la ligne de commande du nœud primaire à l'aide de SSH.
    2. Utilisez un éditeur de texte pour ouvrir le fichier `hue.ini`. Par exemple :

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```
    3. Recherchez le terme `appblacklist` et modifiez la ligne comme suit :

       ```
       appblacklist = search
       ```
    4. Enregistrez vos modifications et redémarrez Hue comme indiqué dans l'exemple suivant :

       ```
       sudo stop hue; sudo start hue
       ```
- Tez
  - Ce problème a été corrigé dans Amazon EMR 5.22.0.

Lorsque vous vous connectez à l'interface utilisateur de Tez à l'adresse http : //:8080/tez-ui *MasterDNS* via une connexion SSH au nœud principal du cluster, l'erreur « L'opération de l'adaptateur a échoué - Le serveur Timeline (ATS) est hors de portée. Soit il est en panne, soit CORS n'est pas activé » apparaît, ou les tâches affichent de manière inattendue « N/A ».

Cela est dû au fait que l'interface utilisateur de Tez envoie des demandes à YARN Timeline Server en utilisant `localhost` plutôt que le nom d'hôte du nœud primaire. Pour contourner le problème, un script est disponible pour être exécuté en tant qu'action ou étape d'amorçage. Le script met à jour le nom d'hôte dans le fichier Tez `configs.env`. Pour plus d'informations et pour connaître l'emplacement du script, consultez les Instructions sur l'amorçage.

- Dans les versions 5.19.0, 5.20.0 et 5.21.0 d'Amazon EMR, les étiquettes des nœuds YARN sont stockées dans un répertoire HDFS. Dans certaines situations, cela entraîne des retards dans le démarrage des nœuds principaux, ce qui provoque le dépassement du délai du cluster et l'échec du lancement. À partir d'Amazon EMR 5.22.0, ce problème est résolu. Les étiquettes des nœuds YARN sont stockées sur le disque local de chaque nœud du cluster, évitant ainsi toute dépendance vis-à-vis du HDFS.
- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.21.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.7.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.5.1	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.11.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.2.0	Connecteur S3Select EMR
emrfs	2.30.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.7.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».

Composant	Version	Description
hadoop-hdfs-datanode	2.8.5-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.5-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	1.4.8	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.8	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.8	HBase client en ligne de commande.
hbase-rest-server	1.4.8	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.8	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.4-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.4-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.4-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.4-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.4-amzn-0	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.4-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.4-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.3.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	0.9.4	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.3.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda

Composant	Version	Description
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.215	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.215	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.0	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.4.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.12.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.1	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.13	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.13	ZooKeeper client en ligne de commande.

Classifications des configurations 5.21.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.21.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.

Classifications	Description
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.

Classifications	Description
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l'MapReduce application.

Classifications	Description
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.

Classifications	Description
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.

Classifications	Description
sqoop-site	Modifiez les valeurs dans le fichier sqoop-sit e.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.20.1

Versions d’application 5.20.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR

- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.20.1	emr-5.20.0	emr-5.19.1	emr-5.19.0
AWS SDK pour Java	1,11,461	1,11,461	1,1,433	1,1,433
Python	2,7, 3,6	2,7, 3,6	2,7, 3,4	2,7, 3,4
Scala	2,1,12	2,1,12	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.2	1.6.2	1.6.1	1.6.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.7	1.4.7
HCatalog	2.3.4	2.3.4	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4

	emr-5.20.1	emr-5.20.0	emr-5.19.1	emr-5.19.0
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.3.1	1.3.1	1.3.0	1.3.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.214	0.214	0.212	0.212
Spark	2.4.0	2.4.0	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.11.0	1.11.0
Tez	0.9.1	0.9.1	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notes de mise à jour 5.20.1

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.20.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.7.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.5.1	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.2.0	Connecteur S3Select EMR
emrfs	2.29.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.6.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».

Composant	Version	Description
hadoop-hdfs-datanode	2.8.5-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.5-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	1.4.8	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.8	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.8	HBase client en ligne de commande.
hbase-rest-server	1.4.8	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.8	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.4-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.4-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.4-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.4-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.4-amzn-0	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.4-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.4-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.3.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	0.9.4	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.3.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda

Composant	Version	Description
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.214	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.214	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.4.0	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.4.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.12.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.1	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.13	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.13	ZooKeeper client en ligne de commande.

Classifications des configurations 5.20.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.20.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.

Classifications	Description
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.

Classifications	Description
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l'MapReduce application.

Classifications	Description
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.

Classifications	Description
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.

Classifications	Description
sqoop-site	Modifiez les valeurs dans le fichier sqoop-sit e.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.20.0

Versions d’application 5.20.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR

- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.20.0	emr-5.19.1	emr-5.19.0	emr-5.18.1
AWS SDK pour Java	1,11,461	1,1,433	1,1,433	1,11,393
Python	2,7, 3,6	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,1,12	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.2	1.6.1	1.6.1	1.6.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.7	1.4.7	1.4.7
HCatalog	2.3.4	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.4
Hive	2.3.4	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.3.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.8.1

	emr-5.20.0	emr-5.19.1	emr-5.19.0	emr-5.18.1
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.3.1	1.3.0	1.3.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.214	0.212	0.212	0.210
Spark	2.4.0	2.3.2	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.11.0	1.11.0	1.9.0
Tez	0.9.1	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.12

Notes de mise à jour 5.20.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.20.0. Les modifications ont été apportées à la version 5.19.0.

Date de parution initiale : 18 décembre 2018

Dernière mise à jour : 22 janvier 2019

Mises à niveau

- Flink 1.6.2
- HBase 1.4.8
- Hive 2.3.4
- Hue 4.3.0
- MXNet 1.3.1
- Presto 0.214
- Spark 2.4.0
- TensorFlow 1,12,0
- Tez 0.9.1
- AWS SDK pour Java 1,11,461

Nouvelles fonctionnalités

- (22 janvier 2019) Kerberos dans Amazon EMR a été amélioré pour prendre en charge l'authentification des principaux à partir d'un KDC externe. Cela permet de centraliser la gestion des principaux, car plusieurs clusters peuvent partager un seul KDC externe. De plus, le KDC externe peut avoir une relation d'approbation inter-domaines avec un domaine Active Directory. Cela permet à tous les clusters d'authentifier les mandataires à partir d'Active Directory. Pour plus d'informations, consultez Utilisation de l'authentification Kerberos dans le Guide de gestion Amazon EMR.

Modifications, améliorations et problèmes résolus

- AMI Amazon Linux par défaut pour Amazon EMR
  - Le package Python3 a été mis à jour de python 3.4 à 3.6.
- Le valideur EMRFS optimisé pour S3
  - Le validateur EMRFS optimisé pour S3 est désormais activé par défaut, ce qui améliore les performances d'écriture. Pour de plus amples informations, veuillez consulter Utilisation d'un valideur EMRFS optimisé pour S3.
- Hive
  - Rétroportage de HIVE-16686.
- Glue avec Spark et Hive

- Dans EMR 5.20.0 ou version ultérieure, l'élagage parallèle des partitions est activé automatiquement pour Spark et Hive lorsque AWS Glue Data Catalog est utilisé comme métastore. Cette modification réduit considérablement le temps de planification des requêtes en exécutant plusieurs requêtes en parallèle pour récupérer des partitions. Le nombre total de segments pouvant être exécutés simultanément est compris entre 1 et 10. La valeur par défaut est 5, ce qui est recommandé. Vous pouvez le modifier en spécifiant la propriété `aws.glue.partition.num.segments` dans la classification de configuration `hive-site`. En cas de limitation, vous pouvez désactiver la fonctionnalité en remplaçant la valeur par 1. Pour en savoir plus, consultez Structure d'un segment AWS Glue.

Problèmes connus

- Hue (corrigée dans la version 5.24.0 d'Amazon EMR)
  - Hue exécuté sur Amazon EMR ne prend pas en charge Solr. À compter de la version 5.20.0 d'Amazon EMR, un problème de configuration incorrect entraîne l'activation de Solr et un message d'erreur inoffensif semblable au suivant s'affiche :

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Pour empêcher l'affichage du message d'erreur de Solr, procédez comme suit :

    1. Connectez-vous à la ligne de commande du nœud primaire à l'aide de SSH.
    2. Utilisez un éditeur de texte pour ouvrir le fichier `hue.ini`. Par exemple :

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```
    3. Recherchez le terme `appblacklist` et modifiez la ligne comme suit :

       ```
       appblacklist = search
       ```
    4. Enregistrez vos modifications et redémarrez Hue comme indiqué dans l'exemple suivant :

       ```
       sudo stop hue; sudo start hue
       ```

- Tez
  - Ce problème a été corrigé dans Amazon EMR 5.22.0.

    Lorsque vous vous connectez à l'interface utilisateur de Tez à l'adresse http : //:8080/tez-ui *MasterDNS* via une connexion SSH au nœud principal du cluster, l'erreur « L'opération de l'adaptateur a échoué - Le serveur Timeline (ATS) est hors de portée. Soit il est en panne, soit CORS n'est pas activé » apparaît, ou les tâches affichent de manière inattendue « N/A ».

    Cela est dû au fait que l'interface utilisateur de Tez envoie des demandes à YARN Timeline Server en utilisant `localhost` plutôt que le nom d'hôte du nœud primaire. Pour contourner le problème, un script est disponible pour être exécuté en tant qu'action ou étape d'amorçage. Le script met à jour le nom d'hôte dans le fichier Tez `configs.env`. Pour plus d'informations et pour connaître l'emplacement du script, consultez les Instructions sur l'amorçage.
- Dans les versions 5.19.0, 5.20.0 et 5.21.0 d'Amazon EMR, les étiquettes des nœuds YARN sont stockées dans un répertoire HDFS. Dans certaines situations, cela entraîne des retards dans le démarrage des nœuds principaux, ce qui provoque le dépassement du délai du cluster et l'échec du lancement. À partir d'Amazon EMR 5.22.0, ce problème est résolu. Les étiquettes des nœuds YARN sont stockées sur le disque local de chaque nœud du cluster, évitant ainsi toute dépendance vis-à-vis du HDFS.
- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Versions des composants 5.20.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.7.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-goodies	2.5.1	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.2.0	Connecteur S3Select EMR
emrfs	2.29.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.6.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.8.5-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.5-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.8.5-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.8	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.8	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.8	HBase client en ligne de commande.
hbase-rest-server	1.4.8	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.8	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.4-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.4-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.4-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.3.4-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.4-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.4-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.4-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.3.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	0.9.4	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.3.1	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.

Composant	Version	Description
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.214	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.214	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques

Composant	Version	Description
spark-client	2.4.0	Clients de ligne de commande Spark.
spark-history-server	2.4.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.4.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.4.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.12.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.9.1	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.0	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.13	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.13	ZooKeeper client en ligne de commande.

Classifications des configurations 5.20.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.20.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.

Classifications	Description
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.

Classifications	Description
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2

Classifications	Description
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.

Classifications	Description
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.

Classifications	Description
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.

Classifications	Description
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.19.1

Versions d’application 5.19.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
AWS SDK pour Java	1,1,433	1,1,433	1,11,393	1,11,393
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.1	1.6.1	1.6.0	1.6.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.7	1.4.7
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.8.1	0.8.1
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.3.0	1.3.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.212	0.212	0.210	0.210
Spark	2.3.2	2.3.2	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.11.0	1.11.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.12	3.4.12

Notes de mise à jour 5.19.1

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.19.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.7.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.5.1	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.1.0	Connecteur S3Select EMR
emrfs	2.28.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.6.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».

Composant	Version	Description
hadoop-hdfs-datanode	2.8.5-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.5-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.5-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	1.4.7	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.7	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.7	HBase client en ligne de commande.
hbase-rest-server	1.4.7	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.7	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.3-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.3-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.3-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.3-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.3-amzn-2	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.3-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.3-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.2.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	0.9.4	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.3.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda

Composant	Version	Description
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.212	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.212	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.2	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.3.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.13	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.13	ZooKeeper client en ligne de commande.

Classifications des configurations 5.19.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.19.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.

Classifications	Description
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.

Classifications	Description
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l'MapReduce application.

Classifications	Description
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.

Classifications	Description
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.

Classifications	Description
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.19.0

Versions d'application 5.19.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR

- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.19.0	emr-5.18.1	emr-5.18.0	emr-5.17.2
AWS SDK pour Java	1,1,433	1,11,393	1,11,393	1,1336
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.1	1.6.0	1.6.0	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.7	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.8.1	0.8.1	0.8.1

	emr-5.19.0	emr-5.18.1	emr-5.18.0	emr-5.17.2
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.3.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.212	0.210	0.210	0.206
Spark	2.3.2	2.3.2	2.3.2	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.11.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.7.3
ZooKeeper	3.4.13	3.4.12	3.4.12	3.4.12

Notes de mise à jour 5.19.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.19.0. Les modifications ont été apportées à la version 5.18.0.

Date de parution initiale : 7 novembre 2018

Date de la dernière mise à jour : 19 novembre 2018

Mises à niveau

- Hadoop 2.8.5
- Flink 1.6.1
- JupyterHub 0,9.4
- MXNet 1.3.0
- Presto 0.212
- TensorFlow 1,11,0
- Zookeeper 3.4.13
- AWS SDK pour Java 1,1,433

Nouvelles fonctionnalités

- (19 novembre 2018) Blocs-notes EMR est un environnement géré basé sur le bloc-notes Jupyter. Il prend en charge les noyaux magiques Spark pour PySpark Spark SQL, Spark R et Scala. Blocs-notes EMR peut être utilisé avec des clusters créés à l'aide de la version 5.18.0 d'Amazon EMR ou d'une version ultérieure. Pour plus d'informations, consultez Utilisation de Blocs-notes EMR dans le Guide de gestion d'Amazon EMR.
- Le validateur EMRFS optimisé pour S3 est disponible lors de l'écriture de fichiers Parquet avec Spark et EMRFS. Ce validateur améliore les performances d'écriture. Pour de plus amples informations, veuillez consulter Utilisation d'un valideur EMRFS optimisé pour S3.

Modifications, améliorations et problèmes résolus

- YARN
  - Modification de la logique qui limite le processus principal de l'application à l'exécution sur les nœuds principaux. Cette fonctionnalité utilise désormais la fonctionnalité et les propriétés des étiquettes de nœuds YARN dans les classifications de configuration `yarn-site` et `capacity-scheduler`. Pour plus d'informations, veuillez consulter https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-plan-instances-guidelines.html#emr-plan-spot-YARN.
- AMI Amazon Linux par défaut pour Amazon EMR
  - `ruby18`, `php56` et `gcc48` ne sont plus installés par défaut. Ils peuvent être installés si vous le souhaitez à l'aide de `yum`.

- Le gem ruby aws-sdk n'est plus installé par défaut. Il peut être installé en utilisant `gem install aws-sdk`, si vous le souhaitez. Des composants spécifiques peuvent également être installés. Par exemple, `gem install aws-sdk-s3`.

Problèmes connus

- Blocs-notes EMR – Dans certains cas, lorsque plusieurs éditeurs de blocs-notes sont ouverts, l'éditeur de bloc-notes peut sembler incapable de se connecter au cluster. Dans ce cas, effacez les cookies du navigateur, puis rouvrez les éditeurs de bloc-notes.
- CloudWatch ContainerPending Mise à l'échelle métrique et automatique — (Corrigé dans la version 5.20.0) Amazon EMR peut émettre une valeur négative pour. `ContainerPending` Si `ContainerPending` est utilisée dans une règle de mise à l'échelle automatique, la mise à l'échelle automatique ne se comporte pas comme prévu. Évitez d'utiliser `ContainerPending` avec la mise à l'échelle automatique.
- Dans les versions 5.19.0, 5.20.0 et 5.21.0 d'Amazon EMR, les étiquettes des nœuds YARN sont stockées dans un répertoire HDFS. Dans certaines situations, cela entraîne des retards dans le démarrage des nœuds principaux, ce qui provoque le dépassement du délai du cluster et l'échec du lancement. À partir d'Amazon EMR 5.22.0, ce problème est résolu. Les étiquettes des nœuds YARN sont stockées sur le disque local de chaque nœud du cluster, évitant ainsi toute dépendance vis-à-vis du HDFS.

Versions des composants 5.19.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.2.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.7.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.5.1	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.1.0	Connecteur S3Select EMR
emrfs	2.28.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.6.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.5-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.5-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.5-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.5-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.5-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.5-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.5-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.8.5-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.5-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.7	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.7	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.7	HBase client en ligne de commande.
hbase-rest-server	1.4.7	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.7	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.3-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.3.3-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.3-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.3-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.3-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.3-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.3-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.2.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	0.9.4	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.3.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.212	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.212	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.2	Clients de ligne de commande Spark.
spark-history-server	2.3.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.11.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.8.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.13	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.13	ZooKeeper client en ligne de commande.

Classifications des configurations 5.19.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.19.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.

Classifications	Description
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.

Classifications	Description
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.

Classifications	Description
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-memory	Modifiez les valeurs dans le fichier memory.properties de Presto.

Classifications	Description
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
presto-connector-tpcds	Modifiez les valeurs dans le fichier tpcds.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.

Classifications	Description
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.18.1

Versions d'application 5.18.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.18.1	emr-5.18.0	emr-5.17.2	emr-5.17.1
AWS SDK pour Java	1,11,393	1,11,393	1,1336	1,1336
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.0	1.6.0	1.5.2	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.6	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-

	emr-5.18.1	emr-5.18.0	emr-5.17.2	emr-5.17.1
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.210	0.210	0.206	0.206
Spark	2.3.2	2.3.2	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notes de mise à jour 5.18.1

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.18.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.1.3	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.6.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-goodies	2.5.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.1.0	Connecteur S3Select EMR
emrfs	2.27.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.6.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.8.4-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.4-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.4-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.4-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.4-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.4-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.4-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.8.4-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.7	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.7	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.7	HBase client en ligne de commande.
hbase-rest-server	1.4.7	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.7	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.3-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.3-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.3-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.3.3-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.3-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.3-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.3-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.2.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	0.8.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.2.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.

Composant	Version	Description
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.210	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.210	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques

Composant	Version	Description
spark-client	2.3.2	Clients de ligne de commande Spark.
spark-history-server	2.3.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.9.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.8.0	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.12	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.12	ZooKeeper client en ligne de commande.

Classifications des configurations 5.18.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.18.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.

Classifications	Description
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.

Classifications	Description
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2

Classifications	Description
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.

Classifications	Description
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.

Classifications	Description
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.18.0

Versions d’application 5.18.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.18.0	emr-5.17.2	emr-5.17.1	emr-5.17.0
AWS SDK pour Java	1,11,393	1,1336	1,1336	1,1336
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.0	1.5.2	1.5.2	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.6	1.4.6	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0

	emr-5.18.0	emr-5.17.2	emr-5.17.1	emr-5.17.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.210	0.206	0.206	0.206
Spark	2.3.2	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notes de mise à jour 5.18.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.18.0. Les modifications ont été apportées à la version 5.17.0.

Date de parution initiale : 24 octobre 2018

Mises à niveau

- Flink 1.6.0
- HBase 1.4.7
- Presto 0.210
- Spark 2.3.2

- Zeppelin 0.8.0

Nouvelles fonctionnalités

- À partir d'Amazon EMR 5.18.0, vous pouvez utiliser le référentiel d'artefacts d'Amazon EMR pour générer le code de votre tâche en fonction des versions exactes des bibliothèques et des dépendances qui sont disponibles avec des versions spécifiques d'Amazon EMR. Pour de plus amples informations, veuillez consulter Vérification des dépendances à l'aide du référentiel d'artefacts d'Amazon EMR.

Modifications, améliorations et problèmes résolus

- Hive
  - Ajout de la prise en charge de S3 Select. Pour de plus amples informations, veuillez consulter Utilisation de S3 Select avec Hive pour améliorer les performances.
- Presto
  - Ajout de la prise en charge de S3 Select Pushdown. Pour de plus amples informations, veuillez consulter Utilisation de S3 Select Pushdown avec Presto pour améliorer les performances.
- Spark
  - La configuration par défaut de log4j pour Spark a été modifiée pour enregistrer les journaux des conteneurs toutes les heures pour les tâches de streaming Spark. Cela permet d'éviter la suppression des journaux pour les tâches de streaming Spark de longue durée.

Versions des composants 5.18.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-`

`component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.1.3	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.6.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.5.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.1.0	Connecteur S3Select EMR
emrfs	2.27.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.6.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.4-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.4-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.4-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.4-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.4-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	2.8.4-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.4-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.4-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.7	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.7	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.7	HBase client en ligne de commande.
hbase-rest-server	1.4.7	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.7	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.3-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.3.3-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.3-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.3-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.3-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.3-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.3-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.2.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	0.8.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
nginx	1.12.1	nginx [engine x] est un serveur HTTP et à proxy inverse

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.2.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.210	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.210	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.2	Clients de ligne de commande Spark.
spark-history-server	2.3.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.9.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.8.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.12	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.12	ZooKeeper client en ligne de commande.

Classifications des configurations 5.18.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.18.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.

Classifications	Description
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.

Classifications	Description
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.

Classifications	Description
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.

Classifications	Description
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.

Classifications	Description
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.17.2

Versions d’application 5.17.2

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.17.2	emr-5.17.1	emr-5.17.0	emr-5.16.1
AWS SDK pour Java	1,1336	1,1336	1,1336	1,1336
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.2	1.5.2	1.5.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.6	1.4.6	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-

	emr-5.17.2	emr-5.17.1	emr-5.17.0	emr-5.16.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.206	0.206	0.203
Spark	2.3.1	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notes de mise à jour 5.17.2

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.17.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.1.3	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.6.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.5.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.0.0	Connecteur S3Select EMR
emrfs	2.26.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.5.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.4-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».

Composant	Version	Description
hadoop-hdfs-datanode	2.8.4-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.4-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.4-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.4-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.4-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.4-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.4-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	1.4.6	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.6	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.4.6	HBase client en ligne de commande.
hbase-rest-server	1.4.6	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.6	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.3-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.3-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.3-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.3-amzn-1	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.3-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.3-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.2.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	0.8.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.2.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.206	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.206	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.1	Clients de ligne de commande Spark.
spark-history-server	2.3.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	2.3.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.9.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.12	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.12	ZooKeeper client en ligne de commande.

Classifications des configurations 5.17.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.17.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop

Classifications	Description
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.

Classifications	Description
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.

Classifications	Description
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-note s Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.

Classifications	Description
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password- authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-en v.sh de Presto.

Classifications	Description
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.

Classifications	Description
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.

Classifications	Description
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.17.1

Versions d’application 5.17.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
AWS SDK pour Java	1,1336	1,1336	1,1336	1,1336
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.2	1.5.0	1.5.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.6	1.4.4	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0,5,0	0,5,0	0,5,0	0,5,0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.206	0.203	0.203
Spark	2.3.1	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notes de mise à jour 5.17.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.17.1. Les modifications ont été apportées à la version 5.17.0.

Date de parution initiale : 18 juillet 2019

Modifications, améliorations et problèmes résolus

- Mise à jour de l'AMI Amazon Linux par défaut pour Amazon EMR afin d'inclure d'importantes mises à jour de sécurité du noyau Linux, notamment le problème de déni de service TCP SACK (AWS-2019-005).

Versions des composants 5.17.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon

EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.1.3	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.6.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.5.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.0.0	Connecteur S3Select EMR
emrfs	2.26.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.5.2	Applications et scripts client de la ligne de commande Apache Flink.

Composant	Version	Description
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.4-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.4-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.4-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.4-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.4-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.4-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.4-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.4-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.6	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.6	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.6	HBase client en ligne de commande.
hbase-rest-server	1.4.6	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.6	Service fournissant un point de terminaison Thrift pour. HBase

Composant	Version	Description
hcatalog-client	2.3.3-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.3-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.3-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.3-amzn-1	Client Hive-hbase.
hive-metastore-server	2.3.3-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.3-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.2.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	0.8.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.2.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.206	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.206	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.1	Clients de ligne de commande Spark.
spark-history-server	2.3.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.9.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.12	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.12	ZooKeeper client en ligne de commande.

Classifications des configurations 5.17.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.17.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.

Classifications	Description
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.

Classifications	Description
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.

Classifications	Description
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-notes Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.

Classifications	Description
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.

Classifications	Description
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.17.0

Versions d’application 5.17.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.17.0	emr-5.16.1	emr-5.16.0	emr-5.15.1
AWS SDK pour Java	1,1336	1,1336	1,1336	1,1333
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.0	1.5.0	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.4	1.4.4	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-

	emr-5.17.0	emr-5.16.1	emr-5.16.0	emr-5.15.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0,5,0	0,5,0	0,5,0	0,4,0
MXNet	1.2.0	1.2.0	1.2.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.203	0.203	0.194
Spark	2.3.1	2.3.1	2.3.1	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notes de mise à jour 5.17.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.17.0. Les modifications ont été apportées à la version 5.16.0.

Date de parution initiale : 30 août 2018

Mises à niveau

- Flink 1.5.2
- HBase 1.4.6
- Presto 0.206

Nouvelles fonctionnalités

- Ajout de la prise en charge de Tensorflow. Pour de plus amples informations, veuillez consulter TensorFlow.

Modifications, améliorations et problèmes résolus

- JupyterHub
  - Ajout de la prise en charge de la persistance des blocs-notes dans Amazon S3. Pour de plus amples informations, veuillez consulter Configuration de la persistance pour les blocs-notes dans Amazon S3.
- Spark
  - Ajout de la prise en charge de S3 Select. Pour de plus amples informations, veuillez consulter Utilisation de S3 Select avec Spark pour améliorer les performances des requêtes.
- Résolution des problèmes liés aux métriques Cloudwatch et à la fonctionnalité de mise à l'échelle automatique dans Amazon EMR version 5.14.0, 5.15.0 ou 5.16.0.

Problèmes connus

- Lorsque vous créez un cluster activé pour Kerberos avec Livy installé, Livy échoue avec un message d'erreur indiquant que l'authentification simple n'est pas activée. Le redémarrage du serveur Livy résout le problème. Pour contourner le problème, ajoutez une étape lors de la création du cluster qui exécute `sudo restart livy-server` sur le nœud primaire.
- Si vous utilisez une AMI Amazon Linux personnalisée basée sur une AMI Amazon Linux dont la date de création est le 11/08/2018, le serveur Oozie ne démarre pas. Si vous utilisez Oozie, créez une AMI personnalisée basée sur un ID d'AMI Amazon Linux avec une date de création différente. Vous pouvez utiliser la AWS CLI commande suivante pour renvoyer une liste d'images IDs pour tous les HVM Amazon Linux AMIs avec une version 2018.03, ainsi que la date de sortie, afin de

pouvoir choisir une AMI Amazon Linux appropriée comme base. MyRegion Remplacez-le par l'identifiant de votre région, tel que us-west-2.

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`]|[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

Versions des composants 5.17.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.1.3	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.6.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.5.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emr-s3-select	1.0.0	Connecteur S3Select EMR
emrfs	2.26.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.5.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.4-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».

Composant	Version	Description
hadoop-hdfs-datanode	2.8.4-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.4-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.4-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.4-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.4-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.4-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.4-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.

Composant	Version	Description
hbase-hmaster	1.4.6	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.6	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.6	HBase client en ligne de commande.
hbase-rest-server	1.4.6	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.6	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.3-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.3-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.3-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.3-amzn-1	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.3-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.3-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.2.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	0.8.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.2.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.206	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.206	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.1	Clients de ligne de commande Spark.
spark-history-server	2.3.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	2.3.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tensorflow	1.9.0	TensorFlow bibliothèque de logiciels open source pour le calcul numérique de haute performance.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.12	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.12	ZooKeeper client en ligne de commande.

Classifications des configurations 5.17.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.17.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop

Classifications	Description
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.

Classifications	Description
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.

Classifications	Description
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-s3-conf	Configuration de la persistance S3 de bloc-note s Jupyter.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.

Classifications	Description
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password- authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-en v.sh de Presto.

Classifications	Description
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.

Classifications	Description
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.

Classifications	Description
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.16.1

Versions d’application 5.16.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
AWS SDK pour Java	1,1336	1,1336	1,1333	1,1333
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4

	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.0	1.5.0	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.4	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0,5,0	0,5,0	0,4,0	0,4,0
MXNet	1.2.0	1.2.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.13.0	4.13.0

	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.203	0.203	0.194	0.194
Spark	2.3.1	2.3.1	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notes de mise à jour 5.16.1

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.16.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent

généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.1.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.6.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.25.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.5.0	Applications et scripts client de la ligne de commande Apache Flink.

Composant	Version	Description
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.4-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.4-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.4-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.4-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.4-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.4-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.8.4-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.4-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.4-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.4-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.4	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.4	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.4	HBase client en ligne de commande.
hbase-rest-server	1.4.4	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.4	Service fournissant un point de terminaison Thrift pour. HBase

Composant	Version	Description
hcatalog-client	2.3.3-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.3-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.3-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.3-amzn-1	Client Hive-hbase.
hive-metastore-server	2.3.3-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.3-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.2.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	0.8.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.2.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.203	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.203	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.1	Clients de ligne de commande Spark.
spark-history-server	2.3.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.12	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.12	ZooKeeper client en ligne de commande.

Classifications des configurations 5.16.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.16.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.

Classifications	Description
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.

Classifications	Description
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2

Classifications	Description
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l'MapReduce application.

Classifications	Description
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.

Classifications	Description
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.

Classifications	Description
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.

Classifications	Description
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.16.0

Versions d’application 5.16.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
AWS SDK pour Java	1,1336	1,1333	1,1333	1,11,297
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.0	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.4	1.4.2
HCatalog	2.3.3	2.3.3	2.3.3	2.3.2
Hadoop	2.8.4	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0,5,0	0,4,0	0,4,0	0,4,0
MXNet	1.2.0	1.1.0	1.1.0	1.1.0

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	4.3.0
Phoenix	4.14.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.203	0.194	0.194	0.194
Spark	2.3.1	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.10

Notes de mise à jour 5.16.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.16.0. Les modifications ont été apportées à la version 5.15.0.

Date de parution initiale : 19 juillet 2018

Mises à niveau

- Hadoop 2.8.4
- Flink 1.5.0
- Livy 0.5.0
- MXNet 1.2.0

- Phoenix 4.14.0
- Presto 0.203
- Spark 2.3.1
- AWS SDK pour Java 1,1336
- CUDA 9.2
- Pilote JDBC Redshift 1.2.15.1025

Modifications, améliorations et problèmes résolus

- HBase
  - Rétroportage de HBASE-20723.
- Presto
  - Modifications de configuration pour prendre en charge l'authentification LDAP. Pour de plus amples informations, veuillez consulter Utilisation de l'authentification LDAP pour Presto sur Amazon EMR.
- Spark
  - Apache Spark version 2.3.1, disponible à partir de la version 5.16.0 d'Amazon EMR, corrige CVE-2018-8024 et CVE-2018-1334. Nous vous recommandons de migrer les versions antérieures de Spark vers la version 2.3.1 ou ultérieure.

Problèmes connus

- Cette version ne prend pas en charge les types d'instance c1.medium ou m1.small. Les clusters utilisant l'un ou l'autre de ces types d'instances ne démarrent pas. Pour contourner le problème, spécifiez un type d'instance différent ou utilisez une version différente.
- Lorsque vous créez un cluster activé pour Kerberos avec Livy installé, Livy échoue avec un message d'erreur indiquant que l'authentification simple n'est pas activée. Le redémarrage du serveur Livy résout le problème. Pour contourner le problème, ajoutez une étape lors de la création du cluster qui exécute `sudo restart livy-server` sur le nœud primaire.
- Après le redémarrage du nœud principal ou le redémarrage du contrôleur d'instance, les CloudWatch métriques ne seront pas collectées et la fonctionnalité de dimensionnement automatique ne sera pas disponible dans les versions 5.14.0, 5.15.0 ou 5.16.0 d'Amazon EMR. Ce problème est résolu dans Amazon EMR 5.17.0.

Versions des composants 5.16.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.1.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.6.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.25.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.5.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.4-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.4-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.4-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.4-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.4-amzn-0	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.8.4-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.4-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.4-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.4-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.4	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.4	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.4	HBase client en ligne de commande.
hbase-rest-server	1.4.4	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.4.4	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.3-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.3-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.3-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.3-amzn-1	Client Hive-hbase.
hive-metastore-server	2.3.3-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.3-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.2.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
jupyterhub	0.8.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.5.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.2.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.2.88	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,14.0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,14.0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica

Composant	Version	Description
presto-coordinator	0.203	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.203	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.1	Clients de ligne de commande Spark.
spark-history-server	2.3.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.12	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.12	ZooKeeper client en ligne de commande.

Classifications des configurations 5.16.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.16.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.

Classifications	Description
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.

Classifications	Description
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.

Classifications	Description
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.

Classifications	Description
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.

Classifications	Description
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-password-authenticator	Modifiez les valeurs dans le fichier password-authenticator.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.

Classifications	Description
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.

Classifications	Description
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.15.1

Versions d’application 5.15.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR

- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
AWS SDK pour Java	1,1333	1,1333	1,11,297	1,11,297
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.2	1.4.2
HCatalog	2.3.3	2.3.3	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1

	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	1.1.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.10	3.4.10

Notes de mise à jour 5.15.1

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.15.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.24.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.4.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.3-amzn-1	HDFS node-level service for storing blocks.

Composant	Version	Description
hadoop-hdfs-library	2.8.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.4	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.4.4	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.4	HBase client en ligne de commande.
hbase-rest-server	1.4.4	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.4	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.3-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.3-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.3-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.3-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.3-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.3-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.

Composant	Version	Description
hive-server2	2.3.3-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.2.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
jupyterhub	0.8.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	1.1.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.1.85	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur

Composant	Version	Description
phoenix-library	4,13,0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,13,0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.194	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.194	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.0	Clients de ligne de commande Spark.
spark-history-server	2.3.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.0	Bibliothèques Apache Spark requises par les esclaves YARN.

Composant	Version	Description
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.12	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.12	ZooKeeper client en ligne de commande.

Classifications des configurations 5.15.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.15.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.

Classifications	Description
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.

Classifications	Description
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.

Classifications	Description
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_n otebook_config.py de Jupyter Notebook.

Classifications	Description
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.

Classifications	Description
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.

Classifications	Description
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile .properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.p roperties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.pro perties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresq l.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.pr operties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.pro perties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift. properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.prop erties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-def aults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Spark.

Classifications	Description
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.15.0

Versions d'application 5.15.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.15.0	emr-5.14.2	emr-5.14.1	emr-5.14.0
AWS SDK pour Java	1,1333	1,11,297	1,11,297	1,11,297
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.2	1.4.2	1.4.2
HCatalog	2.3.3	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-

	emr-5.15.0	emr-5.14.2	emr-5.14.1	emr-5.14.0
Hue	4.2.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	1.1.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.15.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.15.0. Les modifications ont été apportées à la version 5.14.0.

Date de parution initiale : 21 juin 2018

Mises à niveau

- Mise à niveau HBase vers la version 1.4.4
- Mise à niveau de Hive vers la version 2.3.3
- Mise à niveau de Hue vers la version 4.2.0
- Mise à niveau d'Oozie vers la version 5.0.0
- Mise à niveau de Zookeeper vers la version 3.4.12
- AWS SDK mis à jour vers la version 1.11.333

Modifications, améliorations et problèmes résolus

- Hive
  - Rétroportage de HIVE-18069.
- Hue
  - Hue a été mis à jour pour s'authentifier correctement auprès de Livy lorsque Kerberos est activé. Livy est désormais pris en charge lors de l'utilisation de Kerberos avec Amazon EMR.
- JupyterHub
  - Mis à jour JupyterHub afin qu'Amazon EMR installe les bibliothèques clientes LDAP par défaut.
  - Correction d'une erreur dans le script qui génère des certificats auto-signés.

Problèmes connus

- Cette version ne prend pas en charge les types d'instance c1.medium ou m1.small. Les clusters utilisant l'un ou l'autre de ces types d'instances ne démarrent pas. Pour contourner le problème, spécifiez un type d'instance différent ou utilisez une version différente.
- Après le redémarrage du nœud principal ou le redémarrage du contrôleur d'instance, les CloudWatch métriques ne seront pas collectées et la fonctionnalité de dimensionnement automatique ne sera pas disponible dans les versions 5.14.0, 5.15.0 ou 5.16.0 d'Amazon EMR. Ce problème est résolu dans Amazon EMR 5.17.0.

Versions des composants 5.15.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.24.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.4.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.8.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.4	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.4	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.4	HBase client en ligne de commande.
hbase-rest-server	1.4.4	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.4.4	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.3-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.3-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.3-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.3-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.3-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.3-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.3-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.2.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
jupyterhub	0.8.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.1.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.1.85	Pilotes Nvidia et boîte à outils Cuda
oozie-client	5.0.0	Client de ligne de commande Oozie.
oozie-server	5.0.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,13,0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,13,0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica

Composant	Version	Description
presto-coordinator	0.194	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.194	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.0	Clients de ligne de commande Spark.
spark-history-server	2.3.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.12	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.12	ZooKeeper client en ligne de commande.

Classifications des configurations 5.15.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.15.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.

Classifications	Description
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.

Classifications	Description
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.

Classifications	Description
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.

Classifications	Description
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.

Classifications	Description
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.

Classifications	Description
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.14.2

Versions d’application 5.14.2

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
AWS SDK pour Java	1,11,297	1,11,297	1,11,297	1,11,297
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.2
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	-
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	1.1.0	1.1.0	1.1.0	1.0.0

	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.14.2

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.14.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.23.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.4.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.8.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.2	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.2	HBase client en ligne de commande.
hbase-rest-server	1.4.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.4.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.2-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.2-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.1.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
jupyterhub	0.8.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.1.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.1.85	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,13,0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,13,0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica

Composant	Version	Description
presto-coordinator	0.194	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.194	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.0	Clients de ligne de commande Spark.
spark-history-server	2.3.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.14.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.14.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.

Classifications	Description
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.

Classifications	Description
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.

Classifications	Description
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.

Classifications	Description
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.

Classifications	Description
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.

Classifications	Description
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.14.1

Versions d’application 5.14.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.14.1	emr-5.14.0	emr-5.13.1	emr-5.13.0
AWS SDK pour Java	1,11,297	1,11,297	1,11,297	1,11,297
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.2
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	-	-
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	1.1.0	1.1.0	1.0.0	1.0.0

	emr-5.14.1	emr-5.14.0	emr-5.13.1	emr-5.13.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.14.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.14.1. Les modifications ont été apportées à la version 5.14.0.

Date de parution initiale : 17 octobre 2018

Mise à jour de l'AMI par défaut pour Amazon EMR afin de corriger les vulnérabilités de sécurité potentielles.

Versions des composants 5.14.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont

propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.23.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.4.2	Applications et scripts client de la ligne de commande Apache Flink.

Composant	Version	Description
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.2	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.2	HBase client en ligne de commande.
hbase-rest-server	1.4.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.2	Service fournissant un point de terminaison Thrift pour. HBase

Composant	Version	Description
hcatalog-client	2.3.2-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.2-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.1.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
jupyterhub	0.8.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.1.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.1.85	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,13,0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,13,0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.194	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.194	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.0	Clients de ligne de commande Spark.
spark-history-server	2.3.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.14.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.14.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.

Classifications	Description
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.

Classifications	Description
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2

Classifications	Description
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l'MapReduce application.

Classifications	Description
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.

Classifications	Description
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.

Classifications	Description
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.

Classifications	Description
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.14.0

Versions d’application 5.14.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.14.0	emr-5.13.1	emr-5.13.0	emr-5.12.3
AWS SDK pour Java	1,11,297	1,11,297	1,11,297	1,11,267
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	-	-	-
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	1.1.0	1.0.0	1.0.0	1.0.0

	emr-5.14.0	emr-5.13.1	emr-5.13.0	emr-5.12.3
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.188
Spark	2.3.0	2.3.0	2.3.0	2.2.1
Sqoop	1.4.7	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.14.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.14.0. Les modifications ont été apportées à la version 5.13.0.

Date de parution initiale : 4 juin 2018

Mises à niveau

- Mise à niveau d'Apache Flink vers la version 1.4.2
- Mise à niveau d'Apache MXnet vers la version 1.1.0
- Mise à niveau d'Apache Sqoop vers la version 1.4.7

Nouvelles fonctionnalités

- JupyterHub Support ajouté. Pour de plus amples informations, veuillez consulter JupyterHub.

Modifications, améliorations et problèmes résolus

- EMRFS
  - La chaîne userAgent dans les demandes adressées à Amazon S3 a été mise à jour pour contenir les informations relatives à l'utilisateur et au groupe du principal invoquant. Cela peut être utilisé avec AWS CloudTrail les journaux pour un suivi plus complet des demandes.
- HBase
  - Inclusion de HBASE-20447, qui corrige un problème qui pouvait causer des problèmes de cache, en particulier avec les régions divisées.
- MXnet
  - Ajout des bibliothèques OpenCV.
- Spark
  - Lorsque Spark écrit des fichiers Parquet sur un emplacement Amazon S3 à l'aide d'EMRFS, l' FileOutputCommitter algorithme a été mis à jour pour utiliser la version 2 au lieu de la version 1. Cela réduit le nombre de renommages, ce qui améliore les performances de l'application. Cette modification n'affecte pas :
    - Les applications autres que Spark.
    - Applications qui écrivent sur d'autres systèmes de fichiers, tels que HDFS (qui utilise toujours la version 1 de FileOutputCommitter).
    - Les applications utilisant d'autres formats de sortie, tels que le texte ou le csv, qui utilisent déjà l'écriture directe EMRFS.

Problèmes connus

- JupyterHub
  - L'utilisation de classifications de configuration pour configurer JupyterHub des blocs-notes Jupyter individuels lors de la création d'un cluster n'est pas prise en charge. Modifiez manuellement le fichier jupyterhub_config.py et les fichiers jupyter_notebook_config.py pour chaque utilisateur. Pour de plus amples informations, veuillez consulter Configuration JupyterHub.

- JupyterHub ne démarre pas sur des clusters au sein d'un sous-réseau privé, ce qui entraîne un échec du message`Error: ENOENT: no such file or directory, open '/etc/jupyter/conf/server.crt'` . Ce problème est dû à une erreur dans le script qui génère des certificats auto-signés. Utilisez la solution suivante pour générer des certificats auto-signés. Toutes les commandes sont exécutées lorsque vous êtes connecté au nœud primaire.

  1. Copiez le script de génération de certificat du conteneur vers le nœud primaire :

     ```
     sudo docker cp jupyterhub:/tmp/gen_self_signed_cert.sh ./
     ```

  2. Utilisez un éditeur de texte pour modifier la ligne 23 afin de changer le nom d'hôte public en nom d'hôte local, comme indiqué ci-dessous :

     ```
     local hostname=$(curl -s $EC2_METADATA_SERVICE_URI/local-hostname)
     ```

  3. Exécutez le script pour générer des certificats auto-signés :

     ```
     sudo bash ./gen_self_signed_cert.sh
     ```

  4. Déplacez les fichiers de certificat générés par le script vers le répertoire `/etc/jupyter/conf/` :

     ```
     sudo mv /tmp/server.crt /tmp/server.key /etc/jupyter/conf/
     ```

  Vous pouvez vérifier que `tail` le `jupyter.log` fichier JupyterHub a redémarré et qu'il renvoie un code de réponse 200. Par exemple :

  ```
  tail -f /var/log/jupyter/jupyter.log
  ```

  Vous devriez obtenir une réponse similaire à la suivante :

  ```
  # [I 2018-06-14 18:56:51.356 JupyterHub app:1581] JupyterHub is now running at
   https://:9443/
  # 19:01:51.359 - info: [ConfigProxy] 200 GET /api/routes
  ```

- Après le redémarrage du nœud principal ou le redémarrage du contrôleur d'instance, les CloudWatch métriques ne seront pas collectées et la fonctionnalité de dimensionnement automatique ne sera pas disponible dans les versions 5.14.0, 5.15.0 ou 5.16.0 d'Amazon EMR. Ce problème est résolu dans Amazon EMR 5.17.0.

Versions des composants 5.14.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.23.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.4.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.8.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.2	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.2	HBase client en ligne de commande.
hbase-rest-server	1.4.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.4.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.2-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.2-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.1.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
jupyterhub	0.8.1	Serveur multi-utilisateurs pour blocs-notes Jupyter
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.1.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.1.85	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
opencv	3.4.0	Bibliothèque open source Vision par ordinateur
phoenix-library	4,13,0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,13,0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica

Composant	Version	Description
presto-coordinator	0.194	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.194	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.0	Clients de ligne de commande Spark.
spark-history-server	2.3.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.7	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.14.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.14.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
container-log4j	Modifiez les valeurs dans le fichier container-log4j.properties de Hadoop YARN.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.

Classifications	Description
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.

Classifications	Description
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.

Classifications	Description
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
jupyter-notebook-conf	Modifiez les valeurs dans le fichier jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Modifiez les valeurs dans JupyterHubs le fichier jupyterhub_config.py.
jupyter-sparkmagic-conf	Modifiez les valeurs dans le fichier config.json de Sparkmagic.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.

Classifications	Description
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.

Classifications	Description
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.

Classifications	Description
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.13.1

Versions d’application 5.13.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.13.1	emr-5.13.0	emr-5.12.3	emr-5.12.2
AWS SDK pour Java	1,11,297	1,11,297	1,11,267	1,11,267
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2,11.8	2,11.8	2,11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.0	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0

	emr-5.13.1	emr-5.13.0	emr-5.12.3	emr-5.12.2
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.188	0.188
Spark	2.3.0	2.3.0	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.13.1

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.13.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.22.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.4.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.3-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.3-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.3-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.3-amzn-0	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.8.3-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.3-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.3-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.3-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.2	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.2	HBase client en ligne de commande.
hbase-rest-server	1.4.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.4.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.2-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.2-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.1.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.0.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.1.85	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,13,0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,13,0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.194	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.194	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.0	Clients de ligne de commande Spark.
spark-history-server	2.3.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.13.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.13.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.

Classifications	Description
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.

Classifications	Description
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue

Classifications	Description
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.

Classifications	Description
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.

Classifications	Description
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.13.0

Versions d'application 5.13.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
AWS SDK pour Java	1,11,297	1,11,267	1,11,267	1,11,267
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2,11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.0	1.4.0	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.188	0.188	0.188
Spark	2.3.0	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.13.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.13.0. Les modifications ont été apportées à la version 5.12.0.

Mises à niveau

- Mise à niveau de Spark vers la version 2.3.0
- Mise à niveau HBase vers la version 1.4.2
- Mise à niveau vers de Presto vers la version 0.194
- Mise à niveau AWS SDK pour Java vers la version 1.11.297

Modifications, améliorations et problèmes résolus

- Hive
  - Rétroportage de HIVE-15436. Hive améliorée pour ne APIs renvoyer que les vues.

Problèmes connus

- MXNet ne possède actuellement pas de bibliothèques OpenCV.

Versions des composants 5.13.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.10.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.22.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.4.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.3-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.3-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.3-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.3-amzn-0	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.8.3-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.3-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.3-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.3-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.2	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.2	HBase client en ligne de commande.
hbase-rest-server	1.4.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.4.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.2-amzn-2	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.2-amzn-2	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-2	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-2	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.1.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.0.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.1.85	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,13,0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,13,0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.194	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.194	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
r	3.4.1	Projet R pour les calculs statistiques
spark-client	2.3.0	Clients de ligne de commande Spark.
spark-history-server	2.3.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.3.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.3.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.13.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.13.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.

Classifications	Description
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.

Classifications	Description
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue

Classifications	Description
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.

Classifications	Description
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.

Classifications	Description
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.12.3

Versions d'application 5.12.3

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
AWS SDK pour Java	1,11,267	1,11,267	1,11,267	1,11,267
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.4.0	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.188	0.188
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.12.3

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.12.3

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`*

commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.9.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.21.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.4.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les

Composant	Version	Description
		métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.3-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.3-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.3-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.3-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.3-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.3-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.8.3-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.3-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.0	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.0	HBase client en ligne de commande.
hbase-rest-server	1.4.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.2-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.

Composant	Version	Description
hcatalog-webhcat-server	2.3.2-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-1	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.1.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.0.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.

Composant	Version	Description
nvidia-cuda	9.1.85	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,13,0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,13,0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.188	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.188	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
spark-client	2.2.1	Clients de ligne de commande Spark.
spark-history-server	2.2.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	2.2.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.12.3

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.12.3

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.

Classifications	Description
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.

Classifications	Description
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.12.2

Versions d’application 5.12.2

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR

- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.12.2	emr-5.12.1	emr-5.12.0	emr-5.11.4
AWS SDK pour Java	1,11,267	1,11,267	1,11,267	1,11,238
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.4.0	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-5.12.2	emr-5.12.1	emr-5.12.0	emr-5.11.4
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	1.0.0	1.0.0	1.0.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.188	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.12.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.12.2. Les modifications ont été apportées à la version 5.12.1.

Date de parution initiale : 29 août 2018

Modifications, améliorations et problèmes résolus

- Cette version corrige une vulnérabilité de sécurité potentielle.

Versions des composants 5.12.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.9.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.21.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.4.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.3-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.3-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.3-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.3-amzn-0	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.8.3-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.3-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.8.3-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.3-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.0	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.0	HBase client en ligne de commande.
hbase-rest-server	1.4.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.4.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.2-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.2-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-1	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.1.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.0.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.1.85	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,13,0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,13,0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.188	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.188	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
spark-client	2.2.1	Clients de ligne de commande Spark.
spark-history-server	2.2.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.2.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.12.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.12.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.

Classifications	Description
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.

Classifications	Description
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue

Classifications	Description
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.

Classifications	Description
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra .properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.prop erties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.prope rties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.pro perties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile .properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.p roperties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.pro perties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresq l.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.pr operties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.pro perties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift. properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.prop erties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.

Classifications	Description
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.12.1

Versions d'application 5.12.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
AWS SDK pour Java	1,11,267	1,11,267	1,11,238	1,11,238
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
Flink	1.4.0	1.4.0	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	1.0.0	1.0.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.12.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.12.1. Les modifications ont été apportées à la version 5.12.0.

Date de parution initiale : 29 mars 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités potentielles.

Versions des composants 5.12.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-`

`component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.9.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.21.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.4.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les

Composant	Version	Description
		métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.3-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.3-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.3-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.3-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.3-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.3-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.8.3-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.3-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.0	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.0	HBase client en ligne de commande.
hbase-rest-server	1.4.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.2-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.

Composant	Version	Description
hcatalog-webhcat-server	2.3.2-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-1	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.1.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.0.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.

Composant	Version	Description
nvidia-cuda	9.1.85	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,13,0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,13,0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.188	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.188	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
spark-client	2.2.1	Clients de ligne de commande Spark.
spark-history-server	2.2.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	2.2.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.12.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.12.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.

Classifications	Description
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.

Classifications	Description
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.12.0

Versions d'application 5.12.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR

- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.12.0	emr-5.11.4	emr-5.11.3	emr-5.11.2
AWS SDK pour Java	1,11,267	1,11,238	1,11,238	1,11,238
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-5.12.0	emr-5.11.4	emr-5.11.3	emr-5.11.2
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	1.0.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.12.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.12.0. Les modifications ont été apportées à la version 5.11.1.

Mises à niveau

- AWS SDK pour Java 1.11.238 ⇒ 1.11.267. Pour plus d'informations, consultez le AWS SDK for Java Change Log GitHub on.
- Hadoop 2.7.3 ⇒ 2.8.3. Pour plus d'informations, consultez Versions d'Apache Hadoop.

- Flink 1.3.2 ⇒ 1.4.0. Pour plus d'informations, consultez Annonce de publication d'Apache Flink 1.4.0.
- HBase 1.3.1 ⇒ 1.4.0. Pour plus d'informations, consultez l'annonce HBase de sortie.
- Hue 4.0.1 ⇒ 4.1.0. Pour plus d'informations, veuillez consulter les Notes de mise à jour.
- MxNet 0,12,0 ⇒ 10,0. Pour plus d'informations, consultez la section MXNet Change Log on GitHub.
- Presto 0.187 ⇒ 0.188. Pour plus d'informations, veuillez consulter les Notes de mise à jour.

Modifications, améliorations et problèmes résolus

- Hadoop
  - La propriété `yarn.resourcemanager.decommissioning.timeout` a été remplacée par `yarn.resourcemanager.nodemanager-graceful-decommission-timeout-secs`. Vous pouvez utiliser cette propriété pour personnaliser la réduction de la taille du cluster. Pour plus d'informations, consultez Réduction de la taille des clusters dans le Guide de gestion Amazon EMR.
  - La CLI Hadoop a ajouté l'option `-d` à la commande `cp` (copy), qui spécifie la copie directe. Vous pouvez l'utiliser pour éviter de créer un fichier `.COPYING` intermédiaire, ce qui accélère la copie de données entre Amazon S3. Pour plus d'informations, consultez HADOOP-12384.
- Pig
  - Ajout de la classification de configuration `pig-env`, qui simplifie la configuration des propriétés de l'environnement Pig. Pour de plus amples informations, veuillez consulter Configuration des applications.
- Presto
  - Ajout de la classification de configuration, `presto-connector-redshift` que vous pouvez utiliser pour configurer des valeurs dans le fichier de configuration `redshift.properties` de Presto. Pour plus d'informations, consultez Connecteur Redshift dans la documentation Presto, et Configuration des applications.
  - La prise en charge de Presto pour EMRFS a été ajoutée et constitue la configuration par défaut. Les versions précédentes d'Amazon EMR utilisaient PrestOS3FileSystem, qui était la seule option. Pour de plus amples informations, veuillez consulter Configuration d'EMRFS et de PrestOS3 FileSystem.

> **Note**
>
> Si vous interrogez des données sous-jacentes dans Amazon S3 avec Amazon EMR version 5.12.0, des erreurs Presto peuvent se produire. Cela est dû au fait que Presto ne parvient pas à récupérer les valeurs de classification de configuration depuis `emrfs-site.xml`. Pour contourner le problème, créez un sous-répertoire `emrfs` sous `usr/lib/presto/plugin/hive-hadoop2/` et un lien symbolique dans `usr/lib/presto/plugin/hive-hadoop2/emrfs` vers le fichier `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` existant. Redémarrez ensuite le processus presto-server (`sudo presto-server stop` suivi de `sudo presto-server start`).

- Spark
  - SPARK-22036 rétroporté : la BigDecimal multiplication renvoie parfois la valeur nulle.

Problèmes connus

- MXNet n'inclut pas les bibliothèques OpenCV.
- SparkR n'est pas disponible pour les clusters créés à l'aide d'une AMI personnalisée car R n'est pas installé par défaut sur les nœuds du cluster.

Versions des composants 5.12.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0.1	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.9.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.21.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.4.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.8.3-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.8.3-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.8.3-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.8.3-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.8.3-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.8.3-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.8.3-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.8.3-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.4.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.4.0	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.4.0	HBase client en ligne de commande.
hbase-rest-server	1.4.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.4.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.2-amzn-1	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-1	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.

Composant	Version	Description
hcatalog-webhcat-server	2.3.2-amzn-1	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-1	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-1	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.1.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	1.0.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.

Composant	Version	Description
nvidia-cuda	9.1.85	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,13,0- -1,4 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,13,0- -1,4 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.188	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.188	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
spark-client	2.2.1	Clients de ligne de commande Spark.
spark-history-server	2.2.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	2.2.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.12.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.12.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Modifiez les valeurs dans l'environnement Pig.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.

Classifications	Description
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-redshift	Modifiez les valeurs dans le fichier redshift.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.

Classifications	Description
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.11.4

Versions d'application 5.11.4

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR

- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.11.4	emr-5.11.3	emr-5.11.2	emr-5.11.1
AWS SDK pour Java	1,11,238	1,11,238	1,11,238	1,11,238
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-5.11.4	emr-5.11.3	emr-5.11.2	emr-5.11.1
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.11.4

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.11.4

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.8.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.20.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.3.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-6	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-6	HDFS node-level service for storing blocks.

Composant	Version	Description
hadoop-hdfs-library	2.7.3-amzn-6	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-6	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-6	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-6	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-6	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-6	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.3.1	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.2-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.2-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.

Composant	Version	Description
hive-server2	2.3.2-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.0.1	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	0.12.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.0.176	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.187	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.187	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
spark-client	2.2.1	Clients de ligne de commande Spark.
spark-history-server	2.2.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.2.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.11.4

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.11.4

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.

Classifications	Description
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.

Classifications	Description
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-en v.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.prop erties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.

Classifications	Description
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.11.3

Versions d’application 5.11.3

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
AWS SDK pour Java	1,11,238	1,11,238	1,11,238	1,11,238
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.11.3

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.11.3. Les modifications ont été apportées à la version 5.11.2.

Date de parution initiale : 18 juillet 2019

Modifications, améliorations et problèmes résolus

- Mise à jour de l'AMI Amazon Linux par défaut pour Amazon EMR afin d'inclure d'importantes mises à jour de sécurité du noyau Linux, notamment le problème de déni de service TCP SACK (AWS-2019-005).

Versions des composants 5.11.3

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.8.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.20.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.3.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-6	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-6	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-6	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-6	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-6	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-6	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.7.3-amzn-6	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-6	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.2-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.

Composant	Version	Description
hcatalog-webhcat-server	2.3.2-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.0.1	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	0.12.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.

Composant	Version	Description
nvidia-cuda	9.0.176	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.187	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.187	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
spark-client	2.2.1	Clients de ligne de commande Spark.
spark-history-server	2.2.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	2.2.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.11.3

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.11.3

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.

Classifications	Description
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.

Classifications	Description
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.11.2

Versions d’application 5.11.2

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
AWS SDK pour Java	1,11,238	1,11,238	1,11,238	1,11,221
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.1
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.11.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.11.2. Les modifications ont été apportées à la version 5.11.1.

Date de parution initiale : 29 août 2018

Modifications, améliorations et problèmes résolus

- Cette version corrige une vulnérabilité de sécurité potentielle.

Versions des composants 5.11.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.8.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.20.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.3.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-6	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-6	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-6	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-6	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.7.3-amzn-6	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-6	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-6	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-6	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.2-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.2-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.0.1	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	0.12.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.0.176	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.187	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.

Composant	Version	Description
presto-worker	0.187	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
spark-client	2.2.1	Clients de ligne de commande Spark.
spark-history-server	2.2.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.2.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.11.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.11.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.

Classifications	Description
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.

Classifications	Description
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue

Classifications	Description
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.

Classifications	Description
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.

Classifications	Description
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.

Classifications	Description
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.11.1

Versions d’application 5.11.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
AWS SDK pour Java	1,11,238	1,11,238	1,11,221	1,11,221
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	2,7, 3,4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
HCatalog	2.3.2	2.3.2	2.3.1	2.3.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.1	2.3.1
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.11.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.11.1. Il s'agit des modifications apportées à la version Amazon EMR 5.8.0.

Date de parution initiale : 22 janvier 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités associées à l'exécution spéculative (CVE-2017-5715, CVE-2017-5753 et CVE-2017-5754). Pour de plus amples informations, veuillez consulter https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versions des composants 5.11.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.8.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.20.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.3.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-6	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-6	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-6	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-6	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-6	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-6	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.7.3-amzn-6	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-6	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.2-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.

Composant	Version	Description
hcatalog-webhcat-server	2.3.2-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.0.1	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mxnet	0.12.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.

Composant	Version	Description
nvidia-cuda	9.0.176	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.187	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.187	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
spark-client	2.2.1	Clients de ligne de commande Spark.
spark-history-server	2.2.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	2.2.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.11.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.11.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.

Classifications	Description
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.

Classifications	Description
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.11.0

Versions d’application 5.11.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.11.0	emr-5.10.1	emr-5.10.0	emr-5.9.1
AWS SDK pour Java	1,11,238	1,11,221	1,11,221	1,11,1183
Python	2,7, 3,4	2,7, 3,4	2,7, 3,4	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.1	2.3.1	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.1	2.3.1	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	0.12.0	0.12.0	0.12.0	-

	emr-5.11.0	emr-5.10.1	emr-5.10.0	emr-5.9.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.184
Spark	2.2.1	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.11.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.11.0. Les modifications ont été apportées à la version 5.10.0.

Mises à niveau

- Hive 2.3.2
- Spark 2.2.1
- Kit SDK pour Java 1.11.238

Nouvelles fonctionnalités

- Spark
  - Ajout du paramètre `spark.decommissioning.timeout.threshold`, qui améliore le comportement de mise hors service de Spark lors de l'utilisation d'instances Spot. Pour de plus amples informations, veuillez consulter Configuration du comportement de mise hors service du nœud.
  - Ajout du `aws-sagemaker-spark-sdk` composant à Spark, qui installe Amazon SageMaker Spark et les dépendances associées pour l'intégration de Spark à Amazon SageMaker. Vous pouvez utiliser Amazon SageMaker Spark pour créer des pipelines d'apprentissage automatique (ML) Spark à l'aide d'Amazon SageMaker Stages. Pour plus d'informations, consultez le fichier Readme de SageMaker Spark sur GitHub et son utilisation d'Apache Spark avec Amazon SageMaker dans le manuel Amazon SageMaker Developer Guide.

Problèmes connus

- MXNet n'inclut pas les bibliothèques OpenCV.
- Hive 2.3.2 définit `hive.compute.query.using.stats=true` par défaut. Cela entraîne des requêtes pour obtenir des données à partir de statistiques existantes plutôt que directement à partir des données, ce qui peut être déroutant. Par exemple, si vous avez une table avec `hive.compute.query.using.stats=true` et que vous téléchargez de nouveaux fichiers vers la table `LOCATION`, l'exécution d'une demande `SELECT COUNT(*)` sur la table renvoie le nombre des statistiques, plutôt que de récupérer les lignes ajoutées.

  Pour contourner ce problème, utilisez la commande `ANALYZE TABLE` pour collecter de nouvelles statistiques, ou définissez `hive.compute.query.using.stats=false`. Pour en savoir plus, consultez Statistiques dans Hive dans la documentation Apache Hive.

Versions des composants 5.11.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous

nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
aws-sagemaker-spark-sdk	1.0	Kit de développement logiciel Amazon SageMaker Spark
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.8.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.20.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.3.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst

Composant	Version	Description
		ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-6	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-6	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-6	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-6	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-6	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-6	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-6	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-6	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase

Composant	Version	Description
hcatalog-client	2.3.2-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.2-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.2-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.2-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.2-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.2-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.0.1	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.

Composant	Version	Description
mxnet	0.12.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.0.176	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.187	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.187	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.

Composant	Version	Description
spark-client	2.2.1	Clients de ligne de commande Spark.
spark-history-server	2.2.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.2.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.11.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.11.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase

Classifications	Description
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.

Classifications	Description
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.

Classifications	Description
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.

Classifications	Description
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.

Classifications	Description
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.

Classifications	Description
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.10.1

Versions d’application 5.10.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR

- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
AWS SDK pour Java	1,11,221	1,11,221	1,11,1183	1,11,1183
Python	2,7, 3,4	2,7, 3,4	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.1	2.3.1	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.1	2.3.1	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
JupyterHub	-	-	-	-
Livy	0,4,0	0,4,0	0,4,0	0,4,0
MXNet	0.12.0	0.12.0	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.184	0.184
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.10.1

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.10.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-s3-dist-cp	2.7.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.20.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.3.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-5	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-5	Service au niveau du nœud HDFS pour stocker les blocs.
hadoop-hdfs-library	2.7.3-amzn-5	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	2.7.3-amzn-5	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-5	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-5	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-5	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-5	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-5	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-5	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service pour desservir une ou plusieurs HBase régions.

Composant	Version	Description
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.1-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.

Composant	Version	Description
hive-server2	2.3.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.0.1	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	0.12.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.0.176	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.187	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.187	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
spark-client	2.2.0	Clients de ligne de commande Spark.
spark-history-server	2.2.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.2.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.10.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.10.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.

Classifications	Description
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.

Classifications	Description
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-en v.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.prop erties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.

Classifications	Description
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.10.0

Versions d’application 5.10.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
AWS SDK pour Java	1,11,221	1,11,1183	1,11,1183	1,1,1160
Python	2,7, 3,4	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
HCatalog	2.3.1	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.1	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0,4,0	0,4,0	0,4,0	-
MXNet	0.12.0	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0,16,0
Presto	0.187	0.184	0.184	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.10.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.10.0. Il s'agit des modifications apportées à la version Amazon EMR 5.9.0.

Mises à niveau

- AWS SDK pour Java 1,11,221
- Hive 2.3.1
- Presto 0.187

Nouvelles fonctionnalités

- Ajout de la prise en charge de l'authentification Kerberos. Pour plus d'informations, consultez Utilisation de l'authentification Kerberos dans le Guide de gestion Amazon EMR
- Ajout de la prise en charge des rôles IAM pour les demandes EMRFS. Pour plus d'informations, consultez Configuration des rôles IAM pour les requêtes EMRFS vers Amazon S3 dans le Guide de gestion Amazon EMR.
- Ajout de la prise en charge des types d'instances P2 et P3 basées sur GPU. Pour plus d'informations, consultez les sections Instances Amazon EC2 P2 et Instances Amazon EC2 P3. Les pilotes NVIDIA 384.81 et CUDA 9.0.176 sont installés sur ces types d'instance par défaut.
- Ajout de la prise en charge de Apache MXNet.

Modifications, améliorations et problèmes résolus

- Presto

  - Ajout de la prise en charge de l'utilisation du catalogue de données AWS Glue comme métastore Hive par défaut. Pour plus d'informations, consultez la section Utilisation de Presto avec le catalogue de données AWS Glue.
  - Ajout de la prise en charge des fonctions géospatiales.
  - Ajout de la prise en charge du déversement sur le disque pour les jointures.
  - Ajout de la prise en charge du connecteur Redshift.
- Spark
  - Rétroportage de SPARK-20640, ce qui permet de configurer le délai des appels de procédure distante et les tentatives aléatoires des enregistrements de réorganisation à l'aide des propriétés `spark.shuffle.registration.timeout` et `spark.shuffle.registration.maxAttempts`.
  - SPARK-21549 a été rétroporté, qui corrige une erreur survenue lors de l'écriture personnalisée OutputFormat dans des emplacements autres que le HDFS.
- Rétroportage de Hadoop-13270
- Les bibliothèques Numpy, Scipy et Matplotlib ont été supprimées de l'AMI Amazon EMR de base. Si vous avez besoin de ces bibliothèques pour votre application, elles sont disponibles dans le référentiel d'applications. Vous pouvez utiliser une action d'amorçage pour les installer sur tous les nœuds à l'aide de `yum install`.
- L'image AMI Amazon EMR de base n'inclut plus de packages RPM d'application, de telle manière que les packages RPM ne sont plus présents sur les nœuds du cluster. Custom AMIs et l'AMI de base Amazon EMR font désormais référence au référentiel de packages RPM dans Amazon S3.
- En raison de l'introduction de la facturation à la seconde sur Amazon EC2, le comportement de réduction par défaut est désormais de terminer la tâche à la fin de la tâche plutôt que de terminer à l'heure de l'instance. Pour plus d'informations, consultez Configuration de la diminution de la capacité des clusters.

Problèmes connus

- MXNet n'inclut pas les bibliothèques OpenCV.
- Hive 2.3.1 définit `hive.compute.query.using.stats=true` par défaut. Cela entraîne des requêtes pour obtenir des données à partir de statistiques existantes plutôt que directement à partir des données, ce qui peut être déroutant. Par exemple, si vous avez une table avec `hive.compute.query.using.stats=true` et que vous téléchargez de nouveaux fichiers vers

la table `LOCATION`, l'exécution d'une demande `SELECT COUNT(*)` sur la table renvoie le nombre des statistiques, plutôt que de récupérer les lignes ajoutées.

Pour contourner ce problème, utilisez la commande `ANALYZE TABLE` pour collecter de nouvelles statistiques, ou définissez `hive.compute.query.using.stats=false`. Pour en savoir plus, consultez Statistiques dans Hive dans la documentation Apache Hive.

Versions des composants 5.10.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.5.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-s3-dist-cp	2.7.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.20.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.3.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-5	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-5	Service au niveau du nœud HDFS pour stocker les blocs.
hadoop-hdfs-library	2.7.3-amzn-5	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	2.7.3-amzn-5	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-5	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-5	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-5	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-5	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-5	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-5	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service pour desservir une ou plusieurs HBase régions.

Composant	Version	Description
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.1-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.

Composant	Version	Description
hive-server2	2.3.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.0.1	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mxnet	0.12.0	Bibliothèque flexible, évolutive et efficace pour le deep learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
nvidia-cuda	9.0.176	Pilotes Nvidia et boîte à outils Cuda
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.187	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.187	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
spark-client	2.2.0	Clients de ligne de commande Spark.
spark-history-server	2.2.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.2.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.3	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.10.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.10.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.

Classifications	Description
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.

Classifications	Description
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-en v.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.prop erties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.

Classifications	Description
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.9.1

Versions d’application 5.9.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
AWS SDK pour Java	1,11,1183	1,11,1183	1,1,1160	1,1,1160
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0,4,0	0,4,0	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0,16,0	0,16,0
Presto	0.184	0.184	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.9.1

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.9.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.4.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.7.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.19.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.3.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par

Composant	Version	Description
		le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.3-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.3.0-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.0-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.0.1	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.184	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.184	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
spark-client	2.2.0	Clients de ligne de commande Spark.
spark-history-server	2.2.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.2.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.9.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.9.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.

Classifications	Description
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.

Classifications	Description
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-en v.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.prop erties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.

Classifications	Description
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.9.0

Versions d’application 5.9.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Livy, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
AWS SDK pour Java	1,11,1183	1,1,1160	1,1,1160	1,1,1160
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.1	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0,4,0	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0,16,0	0,16,0	0,16,0
Presto	0.184	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.9.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.9.0. Il s'agit des modifications apportées à la version Amazon EMR 5.8.0.

Date de parution : 5 octobre 2017

Dernière date de mise à jour de la fonctionnalité : 12 octobre 2017

Mises à niveau

- AWS SDK pour Java version 1.11.183
- Flink 1.3.2
- Hue 4.0.1
- Pig 0.17.0
- Presto 0.184

Nouvelles fonctionnalités

- Ajout de la prise en charge de Livy (version 0.4.0-incubating). Pour de plus amples informations, veuillez consulter Apache Livy.
- Ajout de la prise en charge de Hue Notebook pour Spark.
- Ajout de la prise en charge des EC2 instances Amazon de la série i3 (12 octobre 2017).

Modifications, améliorations et problèmes résolus

- Spark

  - Ajout d'un nouvel ensemble de fonctionnalités qui permet de s'assurer que Spark gère plus élégamment la terminaison des nœuds suite à un redimensionnement manuel ou à une demande de stratégie de dimensionnement automatique. Pour de plus amples informations, veuillez consulter Configuration du comportement de mise hors service du nœud.
  - Le protocole SSL est utilisé à la place du 3DES pour le chiffrement en transit du service de transfert par blocs, ce qui améliore les performances lors de l'utilisation de types d' EC2 instances Amazon avec AES-NI.
  - Rétroportage de SPARK-21494.
- Zeppelin
  - Rétroportage de ZEPPELIN-2377.
- HBase
  - Ajout du correctif HBASE-18533, qui autorise des valeurs supplémentaires pour la configuration à l'aide de la classification HBase BucketCache de configuration. `hbase-site`
- Hue
  - Ajout de la prise en charge du catalogue de données AWS Glue pour l'éditeur de requêtes Hive dans Hue.
  - Par défaut, les superutilisateurs de Hue peuvent accéder à tous les fichiers auxquels les rôles IAM Amazon EMR sont autorisés à accéder. Les utilisateurs nouvellement créés n'ont pas automatiquement les autorisations d'accéder à l'explorateur de fichiers Amazon S3 et doivent avoir les autorisations `filebrowser.s3_access` activées pour leur groupe.
- Résolution d'un problème en raison duquel les données JSON sous-jacentes créées à l'aide de AWS Glue Data Catalog étaient inaccessibles.

Problèmes connus

- Le lancement du cluster échoue lorsque toutes les applications sont installées et que la taille du volume racine Amazon EBS par défaut n'est pas modifiée. Pour contourner le problème, utilisez la `aws emr create-cluster` commande du AWS CLI et spécifiez un `--ebs-root-volume-size` paramètre plus grand.
- Hive 2.3.0 définit `hive.compute.query.using.stats=true` par défaut. Cela entraîne des requêtes pour obtenir des données à partir de statistiques existantes plutôt que directement à partir des données, ce qui peut être déroutant. Par exemple, si vous avez une table avec `hive.compute.query.using.stats=true` et que vous téléchargez de nouveaux fichiers vers

la table `LOCATION`, l'exécution d'une demande `SELECT COUNT(*)` sur la table renvoie le nombre des statistiques, plutôt que de récupérer les lignes ajoutées.

Pour contourner ce problème, utilisez la commande `ANALYZE TABLE` pour collecter de nouvelles statistiques, ou définissez `hive.compute.query.using.stats=false`. Pour en savoir plus, consultez Statistiques dans Hive dans la documentation Apache Hive.

Versions des composants 5.9.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.4.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-s3-dist-cp	2.7.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.19.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.3.2	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-4	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	2.7.3-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.3.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.0-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.0-amzn-0	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.3.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	4.0.1	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
livy-server	0.4.0-incubating	Interface REST pour interagir avec Apache Spark
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.184	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.184	Service pour exécuter les éléments d'une requête.
pig-client	0.17.0	Client de ligne de commande Pig.
spark-client	2.2.0	Clients de ligne de commande Spark.
spark-history-server	2.2.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.2.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.9.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.9.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.

Classifications	Description
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
livy-conf	Modifiez les valeurs dans le fichier livy.conf de Livy.
livy-env	Modifiez les valeurs dans l'environnement Livy.
livy-log4j	Modifiez les paramètres Livy log4j.properties.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.

Classifications	Description
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-en v.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.prop erties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.

Classifications	Description
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.8.3

Versions d’application 5.8.3

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
AWS SDK pour Java	1,1,1160	1,1,1160	1,1,1160	1,1,1160
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.8.3

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.8.3

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.4.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.6.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.18.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.3.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par

Composant	Version	Description
		le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.3-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.3.0-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.0-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-1	Client de ligne de commande Pig.
spark-client	2.2.0	Clients de ligne de commande Spark.
spark-history-server	2.2.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.2.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.8.3

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.8.3

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.

Classifications	Description
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.

Classifications	Description
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.

Classifications	Description
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark

Classifications	Description
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.8.2

Versions d’application 5.8.2

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
AWS SDK pour Java	1,1,1160	1,1,1160	1,1,1160	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.1	1.3.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
HCatalog	2.3.0	2.3.0	2.3.0	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.8.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.8.2. Les modifications ont été apportées à la version 5.8.1.

Date de parution initiale : 29 mars 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités potentielles.

Versions des composants 5.8.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.4.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.6.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.18.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.3.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par

Composant	Version	Description
		le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.3-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.3.0-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.0-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-1	Client de ligne de commande Pig.
spark-client	2.2.0	Clients de ligne de commande Spark.
spark-history-server	2.2.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.2.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.8.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.8.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.

Classifications	Description
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.

Classifications	Description
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.

Classifications	Description
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark

Classifications	Description
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.8.1

Versions d’application 5.8.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
AWS SDK pour Java	1,1,1160	1,1,1160	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.0	1.3.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
HCatalog	2.3.0	2.3.0	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.1.1	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.8.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.8.1. Il s'agit des modifications apportées à la version Amazon EMR 5.8.0.

Date de parution initiale : 22 janvier 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités associées à l'exécution spéculative (CVE-2017-5715, CVE-2017-5753 et CVE-2017-5754). Pour de plus amples informations, veuillez consulter https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versions des composants 5.8.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.4.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.6.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.18.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.3.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par

Composant	Version	Description
		le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.3-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.3.0-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.0-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-1	Client de ligne de commande Pig.
spark-client	2.2.0	Clients de ligne de commande Spark.
spark-history-server	2.2.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.2.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.8.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.8.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.

Classifications	Description
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.

Classifications	Description
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.

Classifications	Description
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark

Classifications	Description
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.8.0

Versions d’application 5.8.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
AWS SDK pour Java	1,1,1160	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.0	1.3.0	1.2.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.0

	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
HCatalog	2.3.0	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.9.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.1.1	2.1.1	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.8.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.8.0. Il s'agit des modifications apportées à la version Amazon EMR 5.7.0.

Date de parution initiale : 10 août 2017

Dernière date de mise à jour de la fonctionnalité : 25 septembre 2017

Mises à niveau

- AWS SDK 1.11.160
- Flink 1.3.1
- Hive 2.3.0. Pour plus d'informations, consultez Notes de mise à jour sur le site Apache Hive.
- Spark 2.2.0. Pour plus d'informations, consultez Notes de mise à jour sur le site Apache Spark.

Nouvelles fonctionnalités

- Ajout de la prise en charge de l'affichage de l'historique de l'application (25 septembre 2017). Pour plus d'informations, consultez Affichage de l'historique de l'application dans le Guide de gestion Amazon EMR.

Modifications, améliorations et problèmes résolus

- Intégration avec AWS Glue Data Catalog
  - Hive et Spark SQL ont désormais la possibilité d'utiliser AWS Glue Data Catalog comme magasin de métadonnées Hive. Pour plus d'informations, consultez Utilisation du catalogue de données AWS Glue comme métastore pour Hive et Utiliser le catalogue AWS Glue Data Catalog avec Spark sur Amazon EMR.

- Historique de l'application ajouté aux détails du cluster, ce qui vous permet d'afficher les données historiques des applications YARN et des détails supplémentaires pour les applications Spark. Pour plus d'informations, consultez Affichage de l'historique de l'application dans le Guide de gestion Amazon EMR.
- Oozie
  - Rétroportage de OOZIE-2748.
- Hue
  - Rétroportage de HUE-5859
- HBase
  - Ajout d'un correctif pour exposer l'heure de démarrage du serveur HBase principal via Java Management Extensions (JMX) à l'aide `getMasterInitializedTime` de.
  - Ajout d'un correctif qui améliore la date de début du cluster.

Problèmes connus

- Le lancement du cluster échoue lorsque toutes les applications sont installées et que la taille du volume racine Amazon EBS par défaut n'est pas modifiée. Pour contourner le problème, utilisez la `aws emr create-cluster` commande du AWS CLI et spécifiez un `--ebs-root-volume-size` paramètre plus grand.
- Hive 2.3.0 définit `hive.compute.query.using.stats=true` par défaut. Cela entraîne des requêtes pour obtenir des données à partir de statistiques existantes plutôt que directement à partir des données, ce qui peut être déroutant. Par exemple, si vous avez une table avec `hive.compute.query.using.stats=true` et que vous téléchargez de nouveaux fichiers vers la table `LOCATION`, l'exécution d'une demande `SELECT COUNT(*)` sur la table renvoie le nombre des statistiques, plutôt que de récupérer les lignes ajoutées.

  Pour contourner ce problème, utilisez la commande `ANALYZE TABLE` pour collecter de nouvelles statistiques, ou définissez `hive.compute.query.using.stats=false`. Pour en savoir plus, consultez Statistiques dans Hive dans la documentation Apache Hive.
- Spark – Il existe un problème de fuite du gestionnaire de fichiers avec le démon apppusher lors de l'utilisation de Spark, qui peut apparaître après plusieurs heures ou jours pour une tâche Spark de longue durée. Pour résoudre le problème, connectez-vous au nœud primaire et tapez `sudo /etc/init.d/apppusher stop`. Cette commande arrête le démon apppusher, qu'Amazon EMR redémarre automatiquement.
- Historique de l'application

- Les données d'historique pour les exécuteurs Spark inactifs ne sont pas disponibles.
- L'historique de l'application n'est pas disponible pour les clusters qui utilisent une configuration de sécurité pour permettre le chiffrement à la volée.

Versions des composants 5.8.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.4.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.4.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.4.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.6.0	Application de copie distribuée optimisée pour Amazon S3.

Composant	Version	Description
emrfs	2.18.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.3.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-httpfs-server	2.7.3-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.3.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.3.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.3.0-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.3.0-amzn-0	Client Hive-hbase.
hive-metastore-server	2.3.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.3.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.

Composant	Version	Description
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-1	Client de ligne de commande Pig.

Composant	Version	Description
spark-client	2.2.0	Clients de ligne de commande Spark.
spark-history-server	2.2.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.2.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.2.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.8.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.8.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase

Classifications	Description
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.

Classifications	Description
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.

Classifications	Description
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.

Classifications	Description
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.

Classifications	Description
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.7.1

Versions d’application 5.7.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.0	1.3.0	1.2.1	1.2.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.9.0	4.9.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.1	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.7.1

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.7.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.3.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.5.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.18.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.3.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.1.1-amzn-0	Client Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.

Composant	Version	Description
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.

Composant	Version	Description
spark-client	2.1.1	Clients de ligne de commande Spark.
spark-history-server	2.1.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.1.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.7.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.7.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase

Classifications	Description
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.

Classifications	Description
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.

Classifications	Description
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.

Classifications	Description
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.

Classifications	Description
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.7.0

Versions d’application 5.7.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.0	1.2.1	1.2.1	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
Mahout	0.13.0	0.13.0	0.13.0	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.1	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.7.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.7.0. Il s'agit des modifications apportées à la version Amazon EMR 5.6.0.

Date de parution : 13 juillet 2017

Mises à niveau

- Flink 1.3.0
- Phoenix 4.11.0
- Zeppelin 0.7.2

Nouvelles fonctionnalités

- Possibilité ajoutée de spécifier une AMI Amazon Linux personnalisée lorsque vous créez un cluster. Pour en savoir plus, consultez Utilisation d'une image AMI personnalisée.

Modifications, améliorations et problèmes résolus

- HBase
  - Ajout de la possibilité de configurer des clusters HBase en lecture et réplication. Consultez Utilisation d'un cluster de réplica en lecture.
  - Correctifs de nombreux bogues et améliorations
- Presto – ajout de la possibilité de configurer `node.properties`.
- YARN – ajout de la possibilité de configurer `container-log4j.properties`.
- Sqoop – SQOOP-2880 rétroporté, qui introduit un argument vous permettant de définir le répertoire temporaire de Sqoop.

Versions des composants 5.7.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.3.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.5.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.18.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.3.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par

Composant	Version	Description
		le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.1.1-amzn-0	Client Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,11,0- -1,3 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,11,0- -1,3 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.1.1	Clients de ligne de commande Spark.
spark-history-server	2.1.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.1.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.7.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.7.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.

Classifications	Description
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.

Classifications	Description
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.

Classifications	Description
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark

Classifications	Description
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.6.1

Versions d’application 5.6.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.1	1.2.1	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.6.1

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.6.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.3.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.5.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.17.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.2.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par

Composant	Version	Description
		le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.1.1-amzn-0	Client Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.13.0	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,9,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,9,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.1.1	Clients de ligne de commande Spark.
spark-history-server	2.1.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.1.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.6.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.6.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.

Classifications	Description
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.

Classifications	Description
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.

Classifications	Description
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark

Classifications	Description
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.6.0

Versions d’application 5.6.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.1	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.6.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.6.0. Il s'agit des modifications apportées à la version Amazon EMR 5.5.0.

Date de parution : 5 juin 2017

Mises à niveau

- Flink 1.2.1
- HBase 1.3.1
- Mahout 0.13.0. Il s'agit de la première version de Mahout à prendre en charge Spark 2.x dans les versions 5.0 et ultérieures d'Amazon EMR.
- Spark 2.1.1

Modifications, améliorations et problèmes résolus

- Presto
  - Ajout de la possibilité d'activer une communication SSL/TLS sécurisée entre les nœuds Presto en activant le chiffrement en transit à l'aide d'une configuration de sécurité. Pour plus d'informations, consultez Chiffrement des données en transit.
  - Rétroportage de Presto 7661, ce qui ajoute l'option `VERBOSE` à l'instruction `EXPLAIN ANALYZE` pour transmettre des statistiques de bas niveau plus détaillées sur un plan de la requête.

Versions des composants 5.6.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont

propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.3.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.5.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.17.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.2.1	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst

Composant	Version	Description
		ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.1	HBase client en ligne de commande.
hbase-rest-server	1.3.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.3.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.1.1-amzn-0	Client Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
mahout-client	0.13.0	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,9,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,9,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.1.1	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.1.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.1.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.6.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.6.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase

Classifications	Description
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.

Classifications	Description
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.

Classifications	Description
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.

Classifications	Description
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-env	Modifiez les valeurs dans le fichier presto-env.sh de Presto.
presto-node	Modifiez les valeurs dans le fichier node.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.

Classifications	Description
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.5.4

Versions d’application 5.5.4

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.5.4	emr-5.5.3	emr-5.5.2	emr-5.5.1
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.5.4	emr-5.5.3	emr-5.5.2	emr-5.5.1
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.5.4

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.5.4

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.3.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.5.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.16.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.2.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.0	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	1.3.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.1.1-amzn-0	Client Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.

Composant	Version	Description
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,9,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,9,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.

Composant	Version	Description
spark-client	2.1.0	Clients de ligne de commande Spark.
spark-history-server	2.1.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.1.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.5.4

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.5.4

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase

Classifications	Description
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.

Classifications	Description
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.

Classifications	Description
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.

Classifications	Description
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.

Classifications	Description
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.

Classifications	Description
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.5.3

Versions d’application 5.5.3

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie

	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0

	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notes de mise à jour 5.5.3

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.5.3. Les modifications ont été apportées à la version 5.5.2.

Date de parution initiale : 29 août 2018

Modifications, améliorations et problèmes résolus

- Cette version corrige une vulnérabilité de sécurité potentielle.

Versions des composants 5.5.3

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous

nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.3.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.5.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.16.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.2.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.0	HBase client en ligne de commande.
hbase-rest-server	1.3.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.1.1-amzn-0	Client Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,9,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,9,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.1.0	Clients de ligne de commande Spark.
spark-history-server	2.1.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.0	Moteur d'exécution en mémoire pour YARN.

Composant	Version	Description
spark-yarn-slave	2.1.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.5.3

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.5.3

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-l og4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-l ogging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.

Classifications	Description
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra .properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.prop erties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.prope rties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.pro perties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile .properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.p roperties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.pro perties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresq l.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.pr operties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.pro perties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.prop erties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.

Classifications	Description
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.5.2

Versions d'application 5.5.2

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.5.2	emr-5.5.1	emr-5.5.0	emr-5.4.1
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.5.2	emr-5.5.1	emr-5.5.0	emr-5.4.1
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.170	0.166
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.5.2	emr-5.5.1	emr-5.5.0	emr-5.4.1
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.0
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.9

Notes de mise à jour 5.5.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.5.2. Les modifications ont été apportées à la version 5.5.1.

Date de parution initiale : 29 mars 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités potentielles.

Versions des composants 5.5.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-`

`component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.3.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.5.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.16.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.2.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.0	HBase client en ligne de commande.
hbase-rest-server	1.3.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.1.1-amzn-0	Client Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,9,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,9,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.1.0	Clients de ligne de commande Spark.
spark-history-server	2.1.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.0	Moteur d'exécution en mémoire pour YARN.

Composant	Version	Description
spark-yarn-slave	2.1.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.5.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.5.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-l og4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-l ogging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.

Classifications	Description
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra .properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.prop erties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.prope rties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.pro perties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile .properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.p roperties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.pro perties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresq l.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.pr operties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.pro perties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.prop erties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.

Classifications	Description
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.5.1

Versions d'application 5.5.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.166	0.166
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.0	0.7.0
ZooKeeper	3.4.10	3.4.10	3.4.9	3.4.9

Notes de mise à jour 5.5.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.5.1. Il s'agit des modifications apportées à la version Amazon EMR 5.5.0.

Date de parution initiale : 22 janvier 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités associées à l'exécution spéculative (CVE-2017-5715, CVE-2017-5753 et CVE-2017-5754). Pour de plus amples informations, veuillez consulter https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versions des composants 5.5.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`*

commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.3.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.5.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.16.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.2.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.0	HBase client en ligne de commande.
hbase-rest-server	1.3.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.1.1-amzn-0	Client Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,9,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,9,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.1.0	Clients de ligne de commande Spark.
spark-history-server	2.1.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.0	Moteur d'exécution en mémoire pour YARN.

Composant	Version	Description
spark-yarn-slave	2.1.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.5.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.5.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.

Classifications	Description
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra .properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.prop erties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.prope rties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.pro perties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile .properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.p roperties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.pro perties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresq l.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.pr operties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.pro perties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.prop erties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.

Classifications	Description
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.5.0

Versions d'application 5.5.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.5.0	emr-5.4.1	emr-5.4.0	emr-5.3.2
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.5.0	emr-5.4.1	emr-5.4.0	emr-5.3.2
Flink	1.2.0	1.2.0	1.2.0	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.166	0.166	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.5.0	emr-5.4.1	emr-5.4.0	emr-5.3.2
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.0	0.7.0	0.6.2
ZooKeeper	3.4.10	3.4.9	3.4.9	3.4.9

Notes de mise à jour 5.5.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.5.0. Il s'agit des modifications apportées à la version Amazon EMR 5.4.0.

Date de parution : 26 avril 2017

Mises à niveau

- Hue 3.12
- Presto 0.170
- Zeppelin 0.7.1
- ZooKeeper 3.4.10

Modifications, améliorations et problèmes résolus

- Spark
  - Patch Spark rétroporté (SPARK-20115) Correction DAGScheduler permettant de recalculer tous les blocs de shuffle perdus lorsque le service de shuffle externe n'est pas disponible dans la version 2.1.0 de Spark, incluse dans cette version.
- Flink
  - Flink s'appuie désormais sur Scala 2.11. Si vous utilisez l'API et les bibliothèques Scala, nous vous recommandons d'utiliser Scala 2.11 dans vos projets.

- Résolution d'un problème où les valeurs par défaut `HADOOP_CONF_DIR` et `YARN_CONF_DIR` n'étaient pas définies correctement, si bien que `start-scala-shell.sh` ne fonctionnait pas. Ajout également de la possibilité de définir ces valeurs à l'aide de `env.hadoop.conf.dir` et `env.yarn.conf.dir` dans `/etc/flink/conf/flink-conf.yaml` ou la classification de configuration `flink-conf`.
- Introduction d'une nouvelle commande spécifique à EMR, `flink-scala-shell` comme wrapper pour `start-scala-shell.sh`. Nous vous recommandons d'utiliser cette commande au lieu de `start-scala-shell`. La nouvelle commande simplifie l'exécution. Par exemple, `flink-scala-shell -n 2` démarre un shell Flink Scala avec un parallélisme de 2 tâches.
- Introduction d'une nouvelle commande spécifique à EMR, `flink-yarn-session` comme wrapper pour `yarn-session.sh`. Nous vous recommandons d'utiliser cette commande au lieu de `yarn-session`. La nouvelle commande simplifie l'exécution. Par exemple, `flink-yarn-session -d -n 2` démarre une session Flink de longue durée à l'état détaché avec deux gestionnaires de tâches.
- Résolution du problème (FLINK-6125) Commons httpclient n'est plus nuancé dans Flink 1.2.

- Presto
  - Ajout de la prise en charge de l'authentification LDAP. L'utilisation de LDAP avec Presto sur Amazon EMR nécessite que vous activiez l'accès HTTPS pour le coordinateur Presto (`http-server.https.enabled=true` dans `config.properties`). Pour plus de détails sur la configuration, consultez Authentification LDAP dans la documentation Presto.
  - Ajout de la prise en charge de `SHOW GRANTS`.
- AMI Linux Amazon EMR de base
  - Les versions d'Amazon EMR sont désormais basées sur Amazon Linux 2017.03. Pour plus d'informations, consultez Notes de mise à jour Amazon Linux AMI 2017.03.
  - Suppression de Python 2.6 de l'image Linux de base d'Amazon EMR. Python 2.7 et 3.4 sont installés par défaut. Vous pouvez installer Python 2.6 manuellement si nécessaire.

Versions des composants 5.5.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous

nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.3.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.5.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.16.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.2.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.0	HBase client en ligne de commande.
hbase-rest-server	1.3.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.1.1-amzn-0	Client Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.12.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,9,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,9,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.170	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.170	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.1.0	Clients de ligne de commande Spark.
spark-history-server	2.1.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.0	Moteur d'exécution en mémoire pour YARN.

Composant	Version	Description
spark-yarn-slave	2.1.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.10	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.10	ZooKeeper client en ligne de commande.

Classifications des configurations 5.5.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.5.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-l og4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-l ogging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.

Classifications	Description
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.

Classifications	Description
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.4.1

Versions d'application 5.4.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.4.1	emr-5.4.0	emr-5.3.2	emr-5.3.1
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.4.1	emr-5.4.0	emr-5.3.2	emr-5.3.1
Flink	1.2.0	1.2.0	1.1.4	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.166	0.166	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.4.1	emr-5.4.0	emr-5.3.2	emr-5.3.1
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.0	0.7.0	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notes de mise à jour 5.4.1

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.4.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`*

commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.3.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.15.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.2.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.3.0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.0	HBase client en ligne de commande.
hbase-rest-server	1.3.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.1.1-amzn-0	Client Hive-hbase.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,9,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,9,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.166	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.166	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.1.0	Clients de ligne de commande Spark.
spark-history-server	2.1.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.0	Moteur d'exécution en mémoire pour YARN.

Composant	Version	Description
spark-yarn-slave	2.1.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.0	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client en ligne de commande.

Classifications des configurations 5.4.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.4.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-l og4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-l ogging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.

Classifications	Description
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra .properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.prop erties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.prope rties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.pro perties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile .properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.p roperties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.pro perties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresq l.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.pr operties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.pro perties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.prop erties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.

Classifications	Description
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.4.0

Versions d'application 5.4.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
Flink	1.2.0	1.1.4	1.1.4	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.166	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.0	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notes de mise à jour 5.4.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.4.0. Il s'agit des modifications apportées à la version Amazon EMR 5.3.0.

Date de parution : 8 mars 2017

Mises à niveau

- Mise à niveau vers Flink 1.2.0
- Mise à niveau vers la version HBase 1.3.0
- Mise à niveau vers Phoenix 4.9.0

> **Note**
>
> Si vous effectuez une mise à niveau depuis une version antérieure d'Amazon EMR vers Amazon EMR version 5.4.0 ou supérieure et utilisez une indexation secondaire, mettez à niveau les index locaux comme décrit dans la documentation Apache Phoenix. Amazon EMR supprime les configurations requises de la classification `hbase-site`, mais les index doivent être repeuplés. La mise à niveau en ligne et hors ligne des index est prise en charge. Les mises à niveau sont par défaut en ligne, ce qui signifie que les index sont remplis de nouveau lors de l'initialisation depuis les clients Phoenix en version 4.8.0 ou supérieure. Pour spécifier des mises à niveau hors ligne, définissez la configuration `phoenix.client.localIndexUpgrade` sur false dans la classification `phoenix-site`, puis lancez SSH sur le nœud primaire afin d'exécuter `psql [zookeeper] -1`.

- Mise à niveau vers Presto 0.166
- Mise à niveau vers Zeppelin 0.7.0

Modifications et améliorations

- Ajout de la prise en charge des instances r4. Consultez la section Types d' EC2 instances Amazon.

Versions des composants 5.4.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.3.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.15.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.2.0	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-1	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	2.7.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.3.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.3.0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.3.0	HBase client en ligne de commande.
hbase-rest-server	1.3.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.3.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-hbase	2.1.1-amzn-0	Client Hive-hbase.

Composant	Version	Description
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server2	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.11.0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,9,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,9,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica

Composant	Version	Description
presto-coordinator	0.166	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.166	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.1.0	Clients de ligne de commande Spark.
spark-history-server	2.1.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.1.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.7.0	Portable basé sur le web qui permet l'analyse de données interactives.

Composant	Version	Description
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client en ligne de commande.

Classifications des configurations 5.4.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.4.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.

Classifications	Description
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.

Classifications	Description
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue

Classifications	Description
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.

Classifications	Description
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.

Classifications	Description
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.

Classifications	Description
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.3.2

Versions d’application 5.3.2

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR

- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.3.2	emr-5.3.1	emr-5.3.0	emr-5.2.3
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.4	1.1.4	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-5.3.2	emr-5.3.1	emr-5.3.0	emr-5.2.3
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notes de mise à jour 5.3.2

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.3.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.14.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.1.4	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-1	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	2.7.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.2.3	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.3	HBase client en ligne de commande.
hbase-rest-server	1.2.3	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.3	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.

Composant	Version	Description
hive-server	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.1.0	Clients de ligne de commande Spark.
spark-history-server	2.1.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.1.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.6.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.9	ZooKeeper client en ligne de commande.

Classifications des configurations 5.3.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.3.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop

Classifications	Description
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.

Classifications	Description
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.

Classifications	Description
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.

Classifications	Description
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.

Classifications	Description
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.

Classifications	Description
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.3.1

Versions d’application 5.3.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.3.1	emr-5.3.0	emr-5.2.3	emr-5.2.2
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.4	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.3.1	emr-5.3.0	emr-5.2.3	emr-5.2.2
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notes de mise à jour 5.3.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.3.1. Il s'agit des modifications apportées à la version Amazon EMR 5.3.0.

Date de version : 7 février 2017

Modifications mineures pour rétroporter les correctifs Zeppelin et mettre à jour l'AMI par défaut pour Amazon EMR.

Versions des composants 5.3.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont

propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.14.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.1.4	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst

Composant	Version	Description
		ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.3	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.3	HBase client en ligne de commande.
hbase-rest-server	1.2.3	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.2.3	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.

Composant	Version	Description
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.1.0	Clients de ligne de commande Spark.
spark-history-server	2.1.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	2.1.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.1.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.6.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client en ligne de commande.

Classifications des configurations 5.3.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.3.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.

Classifications	Description
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra .properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.prop erties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.prope rties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.pro perties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile .properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.p roperties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.pro perties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresq l.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.pr operties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.pro perties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.prop erties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.

Classifications	Description
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.3.0

Versions d'application 5.3.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
Flink	1.1.4	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.0.2	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notes de mise à jour 5.3.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.3.0. Il s'agit des modifications apportées à la version Amazon EMR 5.2.1.

Date de parution : 26 janvier 2017

Mises à niveau

- Mise à niveau vers Hive 2.1.1
- Mise à niveau vers Hue 3.11.0
- Mise à niveau vers Spark 2.1.0
- Mise à niveau vers Oozie 4.3.0
- Mise à niveau vers Flink 1.1.4

Modifications et améliorations

- Ajout à Hue d'un correctif qui vous permet d'utiliser le paramètre `interpreters_shown_on_wheel` pour configurer les interpréteurs à afficher en premier sur la roue de sélection Notebook quel que soit leur ordre dans le fichier `hue.ini`.
- Ajout de la classification de configuration, `hive-parquet-logging` que vous pouvez utiliser pour configurer des valeurs dans le fichier `parquet-logging.properties` de Hive.

Versions des composants 5.3.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.14.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.1.4	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.7.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.3	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.3	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	1.2.3	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.3	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.1-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.1-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.1-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.1-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	2.1.1-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.11.0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.52	Serveur de base de données MySQL.
oozie-client	4.3.0	Client de ligne de commande Oozie.
oozie-server	4.3.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.1.0	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.1.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.1.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.1.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.6.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client en ligne de commande.

Classifications des configurations 5.3.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.3.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase

Classifications	Description
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.

Classifications	Description
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-parquet-logging	Modifiez les valeurs dans le fichier parquet-logging.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.

Classifications	Description
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.

Classifications	Description
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.

Classifications	Description
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.

Classifications	Description
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.2.3

Versions d’application 5.2.3

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.157.1	0.157.1	0.157.1	0.152.3
Spark	2.0.2	2.0.2	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.8

Notes de mise à jour 5.2.3

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.2.3

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent

généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.2.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.13.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.1.3	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst

Composant	Version	Description
		ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.3	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.3	HBase client en ligne de commande.
hbase-rest-server	1.2.3	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.2.3	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.0-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	2.1.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.10.0-amzn-0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.

Composant	Version	Description
mysql-server	5.5.52	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.0.2	Clients de ligne de commande Spark.
spark-history-server	2.0.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	2.0.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.0.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'application tez YARN.
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client en ligne de commande.

Classifications des configurations 5.2.3

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.2.3

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-l og4j2.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.

Classifications	Description
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.

Classifications	Description
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.

Classifications	Description
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.2.2

Versions d’application 5.2.2

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3

	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.157.1	0.157.1	0.152.3	0.152.3
Spark	2.0.2	2.0.2	2.0.2	2.0.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.8	3.4.8

Notes de mise à jour 5.2.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.2.2. Il s'agit des modifications apportées à la version Amazon EMR 5.2.1.

Date de parution : 2 mai 2017

Problèmes connus résolus depuis les versions précédentes

- SPARK-194459 a été rétroporté, ce qui résout un problème en cas d'échec de la lecture d'une table ORC contenant des colonnes. char/varchar

Versions des composants 5.2.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.2.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.13.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
flink-client	1.1.3	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par

Composant	Version	Description
		le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.3	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.3	HBase client en ligne de commande.
hbase-rest-server	1.2.3	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.3	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.1.0-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	2.1.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.10.0-amzn-0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.52	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.0.2	Clients de ligne de commande Spark.
spark-history-server	2.0.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.0.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.0.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client en ligne de commande.

Classifications des configurations 5.2.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.2.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.

Classifications	Description
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2

Classifications	Description
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.

Classifications	Description
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.

Classifications	Description
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.

Classifications	Description
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.2.1

Versions d’application 5.2.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR

- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.157.1	0.152.3	0.152.3	0.152.3
Spark	2.0.2	2.0.2	2.0.1	2.0.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.8	3.4.8	3.4.8

Notes de mise à jour 5.2.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.2.1. Il s'agit des modifications apportées à la version Amazon EMR 5.2.0.

Date de parution : 29 décembre 2016

Mises à niveau

- Mise à niveau vers Presto 0.157.1. Pour plus d'informations, consultez Notes de mise à jour Presto dans la documentation Presto.
- Mise à niveau vers Zookeeper 3.4.9. Pour plus d'informations, consultez les notes ZooKeeper de version dans la ZooKeeper documentation d'Apache.

Modifications et améliorations

- Ajout de la prise en charge du type d' EC2 instance Amazon m4.16xlarge dans Amazon EMR version 4.8.3 et versions ultérieures, à l'exception des versions 5.0.0, 5.0.3 et 5.2.0.
- Les versions d'Amazon EMR sont désormais basées sur Amazon Linux 2016.09. Pour de plus amples informations, veuillez consulter https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.
- L'emplacement des chemins de configuration Flink et YARN sont désormais définis par défaut dans `/etc/default/flink`, et vous n'avez pas besoin de définir les variables d'environnement `FLINK_CONF_DIR` et `HADOOP_CONF_DIR` lorsque vous exécutez les scripts de pilote `flink` ou `yarn-session.sh` pour lancer des tâches Flink.
- Ajout du support pour les FlinkKinesisConsumer cours.

Problèmes connus résolus depuis les versions précédentes

- Correction d'un problème dans Hadoop où le ReplicationMonitor thread pouvait rester bloqué pendant longtemps en raison d'une course entre la réplication et la suppression du même fichier dans un grand cluster.
- Correction d'un problème en raison duquel ControlledJob #toString échouait avec une exception de pointeur nul (NPE) lorsque le statut de la tâche n'était pas correctement mis à jour.

Versions des composants 5.2.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous

nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.13.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.1.3	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	2.7.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.3	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.3	HBase client en ligne de commande.
hbase-rest-server	1.2.3	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.3	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.1.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.0-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	2.1.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.10.0-amzn-0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.52	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.0.2	Clients de ligne de commande Spark.
spark-history-server	2.0.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.0.2	Moteur d'exécution en mémoire pour YARN.

Composant	Version	Description
spark-yarn-slave	2.0.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'application tez YARN.
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client en ligne de commande.

Classifications des configurations 5.2.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.2.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-l og4j2.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.

Classifications	Description
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daem on-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.

Classifications	Description
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.

Classifications	Description
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.

Classifications	Description
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.2.0

Versions d’application 5.2.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.2.0	emr-5.1.1	emr-5.1.0	emr-5.0.3
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.2

	emr-5.2.0	emr-5.1.1	emr-5.1.0	emr-5.0.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.152.3	0.152.3	0.152.3	0.152.3
Spark	2.0.2	2.0.1	2.0.1	2.0.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.2.0	emr-5.1.1	emr-5.1.0	emr-5.0.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 5.2.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.2.0. Il s'agit des modifications apportées à la version Amazon EMR 5.1.0.

Date de parution : 21 novembre 2016

Modifications et améliorations

- Ajout du mode de stockage Amazon S3 pour HBase.
- Vous permet de spécifier un emplacement Amazon S3 pour le HBase rootdir. Pour plus d'informations, consultez HBase Amazon S3.

Mises à niveau

- Mise à niveau vers Spark 2.0.2

Problèmes connus résolus depuis les versions précédentes

- Correction d'un problème avec /mnt tout en étant limité à 2 To sur les types d'instance uniquement EBS.
- Correction d'un problème avec les journaux instance-controller et logpusher édités dans leurs fichiers .out correspondants au lieu de leurs fichiers .log configurés log4j, qui permutent toutes les heures. Les fichiers .out n'effectuent pas de permutation, cela remplirait finalement la partition /emr. Ce problème affecte uniquement les types d'instance de virtualisation HVM.

Versions des composants 5.2.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.1.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.12.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
flink-client	1.1.3	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-0	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.7.3-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.3	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.3	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	1.2.3	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.3	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.0-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	2.1.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.10.0-amzn-0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.52	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.152.3	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.152.3	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.0.2	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	2.0.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.0.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.0.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client en ligne de commande.

Classifications des configurations 5.2.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.2.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase	Paramètres sélectionnés par Amazon EMR pour Apache. HBase

Classifications	Description
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.

Classifications	Description
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.

Classifications	Description
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.

Classifications	Description
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.

Classifications	Description
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Classifications	Description
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.1.1

Versions d’application 5.1.1

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
Presto	0.152.3	0.152.3	0.152.3	0.150
Spark	2.0.1	2.0.1	2.0.1	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 5.1.1

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 5.1.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous

nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
emr-ddb	4.1.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.11.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.1.3	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	2.7.3-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.3	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.3	HBase client en ligne de commande.
hbase-rest-server	1.2.3	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.3	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.1.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.0-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	2.1.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.10.0-amzn-0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.52	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.152.3	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.152.3	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.0.1	Clients de ligne de commande Spark.
spark-history-server	2.0.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.0.1	Moteur d'exécution en mémoire pour YARN.

Composant	Version	Description
spark-yarn-slave	2.0.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'application tez YARN.
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client en ligne de commande.

Classifications des configurations 5.1.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.1.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.

Classifications	Description
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.

Classifications	Description
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.

Classifications	Description
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra .properties de Presto.

Classifications	Description
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.

Classifications	Description
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.1.0

Versions d’application 5.1.0

Cette version inclut les applications suivantes : Flink, Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.2	1.2.2	1.2.2

	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.152.3	0.152.3	0.150	0.150
Spark	2.0.1	2.0.1	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 5.1.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.1.0. Il s'agit des modifications apportées à la version Amazon EMR 5.0.3.

Date de parution : 3 novembre 2016

Modifications et améliorations

- Support ajouté pour Flink 1.1.3.
- Presto a été ajouté comme une option dans la section ordinateur portable de Hue.

Mises à niveau

- Mise à niveau vers la version HBase 1.2.3
- Mise à niveau vers Zeppelin 0.6.2

Problèmes connus résolus depuis les versions précédentes

- Correction d'un problème avec des requêtes Tez sur Amazon S3 où les fichiers ORC ne fonctionnaient pas aussi bien qu'avec les versions antérieures d'Amazon EMR 4.x.

Versions des composants 5.1.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon

EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.1.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.11.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
flink-client	1.1.3	Applications et scripts client de la ligne de commande Apache Flink.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	2.7.3-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.3	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.3	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.3	HBase client en ligne de commande.
hbase-rest-server	1.2.3	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.3	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	2.1.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.0-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	2.1.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.10.0-amzn-0	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.52	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.152.3	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.152.3	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.0.1	Clients de ligne de commande Spark.
spark-history-server	2.0.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.0.1	Moteur d'exécution en mémoire pour YARN.

Composant	Version	Description
spark-yarn-slave	2.0.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'application tez YARN.
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.2	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client en ligne de commande.

Classifications des configurations 5.1.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.1.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
flink-conf	Modifiez les paramètres flink-conf.yaml.
flink-log4j	Modifiez les paramètres log4j.properties Flink.
flink-log4j-yarn-session	Modifiez les paramètres du j-yarn-session fichier .properties de Flink log4.
flink-log4j-cli	Modifiez les paramètres de propriétés Flink log4j-cli.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.

Classifications	Description
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.

Classifications	Description
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.

Classifications	Description
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra .properties de Presto.

Classifications	Description
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.

Classifications	Description
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.0.3

Versions d’application 5.0.3

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 5.0.3

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.0.3. Il s'agit des modifications apportées à la version Amazon EMR 5.0.0.

Date de parution : 24 octobre 2016

Mises à niveau

- Mise à niveau vers Hadoop 2.7.3
- Mise à niveau vers Presto 0.152.3, qui comprend la prise en charge de l'interface Web Presto. Utilisez le port 8889 du coordinateur Presto pour accéder à l'interface Web Presto. Pour plus d'informations sur l'interface Web Presto, consultez Interface Web dans la documentation Presto.
- Mise à niveau vers Spark 2.0.1
- Les versions d'Amazon EMR sont désormais basées sur Amazon Linux 2016.09. Pour de plus amples informations, veuillez consulter https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Versions des composants 5.0.3

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
emr-ddb	4.1.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.10.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par

Composant	Version	Description
		le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.3-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.1.0-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	2.1.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.10.0-amzn-0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.52	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.152.3	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.152.3	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.0.1	Clients de ligne de commande Spark.
spark-history-server	2.0.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.0.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.0.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client en ligne de commande.

Classifications des configurations 5.0.3

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.0.3

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.

Classifications	Description
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.

Classifications	Description
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.

Classifications	Description
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.

Classifications	Description
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.0.2

Versions d’application 5.0.2

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 5.0.2

Versions des composants 5.0.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.0.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.9.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.7.2-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.2-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.2-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.2-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.1.0-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	2.1.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.10.0-amzn-0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.46	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.150	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.150	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.0.0	Clients de ligne de commande Spark.
spark-history-server	2.0.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.0.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.0.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.1-SNAPSHOT	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client en ligne de commande.

Classifications des configurations 5.0.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.0.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.

Classifications	Description
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.

Classifications	Description
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.

Classifications	Description
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.

Classifications	Description
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.0.1

Versions d’application 5.0.1

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 5.0.1

Versions des composants 5.0.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.0.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.9.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.7.2-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.2-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.2-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.2-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	2.1.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	2.1.0-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	2.1.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.10.0-amzn-0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.46	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.150	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.150	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.0.0	Clients de ligne de commande Spark.
spark-history-server	2.0.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.0.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	2.0.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.1-SNAPSHOT	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client en ligne de commande.

Classifications des configurations 5.0.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.0.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.

Classifications	Description
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4 j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.

Classifications	Description
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.

Classifications	Description
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.

Classifications	Description
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 5.0.0

Versions d’application 5.0.0

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, Zeppelin, et ZooKeeper.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d’Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 5.0.0

Date de parution : 27 juillet 2016

Mises à niveau

- Mise à niveau vers Hive 2.1
- Mise à niveau vers Presto 0.150
- Mise à niveau vers Spark 2.0
- Mise à niveau vers Hue 3.10.0
- Mise à niveau vers Pig 0.16.0
- Mise à niveau vers Tez 0.8.4
- Mise à niveau vers Zeppelin 0.6.1

Modifications et améliorations

- Amazon EMR prend en charge les dernières versions open source de Hive (version 2.1) et de Pig (version 0.16.0). Si vous utilisiez Hive ou Pig sur Amazon EMR auparavant, cela risque d'avoir un impact sur certains cas d'utilisation. Pour en savoir plus, consultez Hive et Pig.
- Tez est désormais le moteur d'exécution par défaut de Hive et Pig. Pour changer cela, vous devez modifier les valeurs appropriées dans les classifications de configuration `hive-site` et `pig-properties`, respectivement.
- Une fonction améliorée de débogage des étapes a été ajoutée, ce qui vous permet de voir la cause première des échecs des étapes si le service peut déterminer la cause. Pour plus d'informations, consultez Débogage d'étape amélioré dans le Guide de gestion Amazon EMR.
- Les applications qui se terminaient par « -Sandbox » n'ont plus ce suffixe. Cela peut interrompre l'automatisation de certaines tâches, par exemple, si vous utilisez des scripts pour lancer des clusters avec ces applications. Le tableau suivant montre les noms d'application dans Amazon EMR 4.7.2 contre Amazon EMR 5.0.0.

  Modification des noms d'application

Amazon EMR 4.7.2	Amazon EMR 5.0.0
Oozie-Sandbox	Oozie
Presto-Sandbox	Presto
Sqoop-Sandbox	Sqoop
Zeppelin-Sandbox	Zeppelin
ZooKeeper-Bac à sable	ZooKeeper

- Spark est maintenant compilé pour Scala 2.11.
- Java 8 est désormais la JVM par défaut. Toutes les applications s'exécutent avec le runtime Java 8. Aucune modification n'est apportée bytecode cible des applications. La plupart des applications continuent de cibler Java 7.
- Zeppelin inclut désormais des fonctions d'authentification. Pour en savoir plus, consultez Zeppelin.
- Ajout de la prise en charge des configurations de sécurité, qui vous permettent de créer et d'appliquer des options de chiffrement plus facilement. Pour en savoir plus, consultez Chiffrement des données.

Versions des composants 5.0.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.0.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.9.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst

Composant	Version	Description
		ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.2-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.2-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.2-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.2-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service pour desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase

Composant	Version	Description
hcatalog-client	2.1.0-amzn-0	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	2.1.0-amzn-0	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	2.1.0-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	2.1.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	2.1.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.10.0-amzn-0	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.46	Serveur de base de données MySQL.

Composant	Version	Description
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4,7,0- -1,2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4,7,0- -1,2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.150	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.150	Service pour exécuter les éléments d'une requête.
pig-client	0.16.0-amzn-0	Client de ligne de commande Pig.
spark-client	2.0.0	Clients de ligne de commande Spark.
spark-history-server	2.0.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	2.0.0	Moteur d'exécution en mémoire pour YARN.

Composant	Version	Description
spark-yarn-slave	2.0.0	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'application tez YARN.
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.1-SNAPSHOT	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client en ligne de commande.

Classifications des configurations 5.0.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-5.0.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j2	Modifiez les valeurs dans HCatalog HCat log4j2.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat du Web.
hive-beeline-log4j2	Modifiez les valeurs dans le fichier beeline-log4j2.properties de Hive.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j2	Modifiez les valeurs dans le fichier hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Modifiez les valeurs dans le fichier llap-daemon-log 4j2.properties de Hive.
hive-log4j2	Modifiez les valeurs dans le fichier hive-log4j2.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2

Classifications	Description
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.

Classifications	Description
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.

Classifications	Description
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Spark
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.

Classifications	Description
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Versions Amazon EMR 4.x

Cette section contient les versions des applications, les notes de mise à jour, les versions des composants et les classifications de configuration disponibles dans chaque version Amazon EMR 4.x.

Lorsque vous lancez un cluster, vous pouvez choisir parmi plusieurs versions d'Amazon EMR. Vous pouvez ainsi tester et utiliser des versions d'applications qui répondent à vos exigences de compatibilité. Vous spécifiez le numéro de version avec l'étiquette de version. Les étiquettes de version sont sous la forme `emr-x.x.x`. Par exemple, `emr-7.9.0`.

Les nouvelles versions d'Amazon EMR sont mises à disposition dans différentes régions sur une période de plusieurs jours, en commençant par la première région à la date de sortie initiale. Il est possible que la dernière version ne soit pas disponible dans votre région pendant cette période.

Pour un tableau complet des versions des applications dans chaque version 4.x d'Amazon EMR, consultez Versions des applications dans les versions 4.x d'Amazon EMR.

Rubriques

- Versions des applications dans les versions 4.x d'Amazon EMR
- Différences entre les versions Amazon EMR 4.x
- Amazon EMR version 4.9.6
- Amazon EMR version 4.9.5
- Amazon EMR version 4.9.4
- Amazon EMR version 4.9.3
- Amazon EMR version 4.9.2
- Amazon EMR version 4.9.1
- Amazon EMR version 4.8.5
- Amazon EMR version 4.8.4
- Amazon EMR version 4.8.3
- Amazon EMR version 4.8.2
- Amazon EMR version 4.8.1
- Amazon EMR version 4.8.0
- Amazon EMR version 4.7.4
- Amazon EMR version 4.7.3
- Amazon EMR version 4.7.2
- Amazon EMR version 4.7.1
- Amazon EMR version 4.7.0
- Amazon EMR version 4.6.1
- Amazon EMR version 4.6.0
- Amazon EMR version 4.5.0
- Amazon EMR version 4.4.0
- Amazon EMR version 4.3.0
- Amazon EMR version 4.2.0
- Amazon EMR version 4.1.0
- Amazon EMR version 4.0.0

Versions des applications dans les versions 4.x d'Amazon EMR

Pour obtenir un tableau complet répertoriant les versions d'applications disponibles dans chaque version d'Amazon EMR 4.x, ouvrez Versions des applications dans les versions 4.x d'Amazon EMR dans votre navigateur.

Différences entre les versions Amazon EMR 4.x

La documentation relative aux fonctionnalités d'Amazon EMR dans le Guide de gestion Amazon EMR précise la version d'Amazon EMR dans laquelle une fonctionnalité est devenue disponible, ainsi que les différences applicables entre les fonctionnalités d'Amazon EMR qui remontent à la version 4.0.0.

À partir de la version Amazon EMR 5.0.0, certaines applications ont subi une mise à niveau de version importante qui a modifié les modalités d'installation ou opérationnelles, et d'autres applications sont passées du statut d'application sandbox au statut d'application native. Chaque rubrique de cette section présente les différences significatives spécifiques aux applications lorsque vous utilisez les versions Amazon EMR 4.x.

Rubriques

- Applications de l'environnement de test (sandbox)
- Considérations relatives à l'utilisation de Hive sur Amazon EMR 4.x
- Considérations relatives à l'utilisation de Pig sur Amazon EMR 4.x

Applications de l'environnement de test (sandbox)

Lorsque vous utilisez les versions 4.x d'Amazon EMR, certaines applications sont considérées comme des applications d'environnement de test (sandbox). Les applications de l'environnement de test (sandbox) sont des versions antérieures de l'application que nous avons mises à disposition au moment du lancement de la première version d'Amazon EMR en raison de la demande. Vous pouvez utiliser la console ou l'API pour qu'Amazon EMR installe les applications sandbox de la même manière que les applications natives, mais le support et la documentation des applications sandbox sont limités. AWS CLI Les applications sandbox deviennent des applications natives et entièrement prises en charge dans la version Amazon EMR 5.0.0 et versions ultérieures. Voici des applications sandbox dans les versions Amazon EMR 4.x :

- Oozie
- Presto

- Sqoop
- Zeppelin
- ZooKeeper

Lorsque vous installez des applications sandbox, le nom des applications porte le suffixe `-sandbox`. Par exemple, pour installer la version sandbox de*`Presto`*, utilisez`Presto-sandbox`. L'installation de cette application peut prendre plus de temps que celle d'une application entièrement prise en charge. Les numéros de version répertoriés pour chaque application dans cette section correspondent à la version de communauté de l'application.

Oozie (versions de l'environnement de test)

Oozie est disponible sous la forme d'une application d'environnement de test (sandbox) à partir de la version Amazon EMR 4.1.0.

Les exemples Oozie ne sont pas installés par défaut avec les versions sandbox. Pour installer les exemples, accédez au nœud principal via SSH et exécutez `install-oozie-examples`.

Informations sur la version Oozie-Sandbox

Étiquette de version Amazon EMR	Oozie – Version d'environnement de test (sandbox)	Composants installés avec Oozie-Sandbox
emr-4.9.6	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.5	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-ma

Étiquette de version Amazon EMR	Oozie – Version d'environnement de test (sandbox)	Composants installés avec Oozie-Sandbox
		pred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Étiquette de version Amazon EMR	Oozie – Version d'environnement de test (sandbox)	Composants installés avec Oozie-Sandbox
emr-4.9.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Étiquette de version Amazon EMR	Oozie – Version d'environnement de test (sandbox)	Composants installés avec Oozie-Sandbox
emr-4.8.5	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.8.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Étiquette de version Amazon EMR	Oozie – Version d'environnement de test (sandbox)	Composants installés avec Oozie-Sandbox
emr-4.8.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.8.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Étiquette de version Amazon EMR	Oozie – Version d'environnement de test (sandbox)	Composants installés avec Oozie-Sandbox
emr-4.8.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.8.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Étiquette de version Amazon EMR	Oozie – Version d'environnement de test (sandbox)	Composants installés avec Oozie-Sandbox
emr-4.7.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.7.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Étiquette de version Amazon EMR	Oozie – Version d'environnement de test (sandbox)	Composants installés avec Oozie-Sandbox
emr-4.7.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.7.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Étiquette de version Amazon EMR	Oozie – Version d'environnement de test (sandbox)	Composants installés avec Oozie-Sandbox
emr-4.7.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.6.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.6.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Étiquette de version Amazon EMR	Oozie – Version d'environnement de test (sandbox)	Composants installés avec Oozie-Sandbox
emr-4.5.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.4.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.3.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Étiquette de version Amazon EMR	Oozie – Version d'environnement de test (sandbox)	Composants installés avec Oozie-Sandbox
emr-4.2.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.1.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Presto (versions de l'environnement de test)

Presto est disponible sous la forme d'une application d'environnement de test (sandbox) à partir de la version Amazon EMR 4.1.0.

Informations sur la version Presto-Sandbox

Étiquette de version Amazon EMR	Presto – Version d'environnement de test (sandbox)	Composants installés avec Presto-Sandbox
emr-4.9.6	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Étiquette de version Amazon EMR	Presto – Version d'environnement de test (sandbox)	Composants installés avec Presto-Sandbox
		hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.5	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.4	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Presto – Version d'environnement de test (sandbox)	Composants installés avec Presto-Sandbox
emr-4.9.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Presto – Version d'environnement de test (sandbox)	Composants installés avec Presto-Sandbox
emr-4.8.5	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.4	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Presto – Version d'environnement de test (sandbox)	Composants installés avec Presto-Sandbox
emr-4.8.2	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.1	0.151	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.0	0.151	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Presto – Version d'environnement de test (sandbox)	Composants installés avec Presto-Sandbox
emr-4.7.4	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.3	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.2	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Presto – Version d'environnement de test (sandbox)	Composants installés avec Presto-Sandbox
emr-4.7.1	0.147	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.0	0.147	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.6.1	0.143	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Presto – Version d'environnement de test (sandbox)	Composants installés avec Presto-Sandbox
emr-4.6.0	0.143	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.5.0	0.140	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.4.0	0.136	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Presto – Version d'environnement de test (sandbox)	Composants installés avec Presto-Sandbox
emr-4.3.0	0.130	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.2.0	0.125	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.1.0	0.119	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Sqoop (versions de l'environnement de test)

Sqoop est disponible sous la forme d'une application d'environnement de test (sandbox) à partir de la version Amazon EMR 4.4.0.

Informations sur la version Sqoop-Sandbox

Étiquette de version Amazon EMR	Sqoop – Version d'environnement de test (sandbox)	Composants installés avec Sqoop-Sandbox
emr-4.9.6	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.5	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Sqoop – Version d'environnement de test (sandbox)	Composants installés avec Sqoop-Sandbox
emr-4.9.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Sqoop – Version d'environnement de test (sandbox)	Composants installés avec Sqoop-Sandbox
emr-4.8.5	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Sqoop – Version d'environnement de test (sandbox)	Composants installés avec Sqoop-Sandbox
emr-4.8.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Sqoop – Version d'environnement de test (sandbox)	Composants installés avec Sqoop-Sandbox
emr-4.7.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Sqoop – Version d'environnement de test (sandbox)	Composants installés avec Sqoop-Sandbox
emr-4.7.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.6.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Sqoop – Version d'environnement de test (sandbox)	Composants installés avec Sqoop-Sandbox
emr-4.6.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.5.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, sqoop-client
emr-4.4.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, sqoop-client

Zeppelin (versions de l'environnement de test)

Zeppelin est disponible sous la forme d'une application d'environnement de test (sandbox) à partir de la version Amazon EMR 4.1.0.

Informations sur la version Zeppelin-Sandbox

Étiquette de version Amazon EMR	Version Zeppelin-Sandbox	Composants installés avec Zeppelin-Sandbox
emr-4.9.6	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.5	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.4	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-

Étiquette de version Amazon EMR	Version Zeppelin-Sandbox	Composants installés avec Zeppelin-Sandbox
		on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin-Sandbox	Composants installés avec Zeppelin-Sandbox
emr-4.9.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.5	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin-Sandbox	Composants installés avec Zeppelin-Sandbox
emr-4.8.4	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin-Sandbox	Composants installés avec Zeppelin-Sandbox
emr-4.8.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin-Sandbox	Composants installés avec Zeppelin-Sandbox
emr-4.8.0	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.4	0,5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin-Sandbox	Composants installés avec Zeppelin-Sandbox
emr-4.7.3	0,5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.2	0,5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin-Sandbox	Composants installés avec Zeppelin-Sandbox
emr-4.7.1	0,5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.0	0,5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.6.1	0,5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin-Sandbox	Composants installés avec Zeppelin-Sandbox
emr-4.6.0	0,5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-ya rn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.5.0	0,5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-ya rn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.4.0	0,5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-ya rn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin-Sandbox	Composants installés avec Zeppelin-Sandbox
emr-4.3.0	0,5,5	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.2.0	0,5,5	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.1.0	0.6.0-SNAPSHOT	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

ZooKeeper (versions bac à sable)

Zookeeper est disponible sous la forme d'une application d'environnement de test (sandbox) à partir de la version Amazon EMR 4.6.0.

ZooKeeper-Informations sur la version du bac à sable

Étiquette de version Amazon EMR	ZooKeeper-Version bac à sable	Composants installés avec ZooKeeper -Sandbox
emr-4.9.6	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.5	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.4	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Étiquette de version Amazon EMR	ZooKeeper-Version bac à sable	Composants installés avec ZooKeeper -Sandbox
		cemanager, zookeeper-client, zookeeper-server
emr-4.9.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper-Version bac à sable	Composants installés avec ZooKeeper -Sandbox
emr-4.8.5	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.4	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper-Version bac à sable	Composants installés avec ZooKeeper -Sandbox
emr-4.8.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper-Version bac à sable	Composants installés avec ZooKeeper -Sandbox
emr-4.7.4	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.3	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper-Version bac à sable	Composants installés avec ZooKeeper -Sandbox
emr-4.7.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.6.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper-Version bac à sable	Composants installés avec ZooKeeper -Sandbox
emr-4.6.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Considérations relatives à l'utilisation de Hive sur Amazon EMR 4.x

Cette section couvre les différences à prendre en compte lors de l'utilisation de la version Hive 1.0.0 sur les versions 4.x d'Amazon EMR, en comparaison avec Hive 2.x sur les versions 5.x d'Amazon EMR.

Les transactions ACID ne sont pas prises en charge.

Hive sur les versions 4.x d'Amazon EMR ne prend pas en charge les transactions ACID avec les données Hive stockées dans Amazon S3 lorsque vous utilisez les versions 4.x. Si vous essayez de créer une table transactionnelle dans Amazon S3, cela entraîne une exception.

Lecture et écriture de tables dans Amazon S3

Hive sur les versions 4.x d'Amazon EMR peut écrire directement dans Amazon S3 sans utiliser de fichiers temporaires. Cela améliore les performances, mais par voie de conséquence, vous ne pouvez pas lire et écrire dans la même table dans Amazon S3 via la même instruction Hive. Une solution de contournement consiste à créer et utiliser une table temporaire dans HDFS.

L'exemple suivant montre comment utiliser plusieurs instructions Hive pour mettre à jour une table dans Amazon S3. Les instructions créent une table temporaire nommée `tmp` dans HDFS en se basant sur une table dans Amazon S3 nommée `my_s3_table`. La table dans Amazon S3 est alors mise à jour avec le contenu de la table temporaire.

```
CREATE TEMPORARY TABLE tmp LIKE my_s3_table;
INSERT OVERWRITE TABLE tmp SELECT ....;
```

```
INSERT OVERWRITE TABLE my_s3_table SELECT * FROM tmp;
```

Log4j par rapport à Log4j 2

Hive sur les versions Amazon EMR 4.x utilise Log4j. À partir de la version 5.0.0, Log4j 2 est la configuration par défaut. Ces versions peuvent nécessiter différentes configurations de journalisation. Pour plus de détails, consultez Apache Log4j 2.

MapReduce est le moteur d'exécution par défaut

Les versions 4.x de Hive on Amazon EMR sont MapReduce utilisées comme moteur d'exécution par défaut. À partir de la version Amazon EMR 5.0.0, Tez est la valeur par défaut, ce qui offre des performances améliorées pour la plupart des flux de travail.

Autorisation Hive

Hive sur les versions Amazon EMR 4.x prend en charge l'autorisation Hive pour HDFS, mais pas pour EMRFS et Amazon S3. Les clusters Amazon EMR s'exécutent avec l'autorisation désactivée par défaut.

Comportement de fusion de fichiers Hive avec Amazon S3

Hive sur les versions 4.x d'Amazon EMR fusionne de petits fichiers à la fin d'une tâche de mappage uniquement si `hive.merge.mapfiles` est défini sur `true`. Une fusion est déclenchée uniquement si la taille de sortie moyenne de la tâche est inférieure au paramètre `hive.merge.smallfiles.avgsize`. Le comportement d'Amazon EMR Hive est identique si le chemin de la sortie finale est sur HDFS. Cependant, si le chemin de sortie se trouve dans Amazon S3, le paramètre `hive.merge.smallfiles.avgsize` est ignoré. Dans ce cas, la tâche de fusion est toujours déclenchée si `hive.merge.mapfiles` est défini sur `true`.

Considérations relatives à l'utilisation de Pig sur Amazon EMR 4.x

La version Pig 0.14.0 est installée sur les clusters créés avec les versions Amazon EMR 4.x. Pig a été mis à niveau à la version 0.16.0 dans Amazon EMR 5.0.0. Les différences importantes sont discutées ci-dessous.

Moteur d'exécution par défaut différent

La version 0.14.0 de Pig sur les versions 4.x d'Amazon EMR est utilisée MapReduce comme moteur d'exécution par défaut. Pig 0.16.0 et les versions ultérieures utilisent Apache Tez. Vous pouvez

définir explicitement `exectype=mapreduce` la classification `pig-properties` de configuration à utiliser MapReduce.

Fonctions définies par l'utilisateur Dropped Pig () UDFs

Les versions personnalisées UDFs disponibles dans les versions 4.x de Pig on Amazon EMR ont été supprimées à partir de Pig 0.16.0. La plupart d'entre eux UDFs ont des fonctions équivalentes que vous pouvez utiliser à la place. Le tableau suivant répertorie les fonctions supprimées UDFs et équivalentes. Pour plus d'informations, consultez Fonctions intégrées sur le site Apache Pig.

Fonctions définies par l'utilisateur abandonnées	Fonction équivalente
FORMAT_DT (formatdt, date)	GetHour(date), GetMinute (date), GetMonth (date), GetSecond (date), GetWeek (date), GetYear (date), GetDay (date)
EXTRACT(chaîne, modèle)	REGEX_EXTRACT_ALL(chaîne, modèle)
REPLACE(chaîne, modèle, remplacement)	REPLACE(chaîne, modèle, remplacement)
DATE_TIME()	ToDate()
DURATION(dt, dt2)	WeeksBetween(dt, dt2), YearsBetween (dt, dt2), SecondsBetween (dt, dt2), (dt, dt2), MonthsBetween (dt, dt2), MinutesBetween (dt, dt2) HoursBetween
EXTRACT_DT(format, date)	GetHour(date), GetMinute (date), GetMonth (date), GetSecond (date), GetWeek (date), GetYear (date), GetDay (date)
OFFSET_DT(date, durée)	AddDuration(date, durée), SubtractDuration (date, durée)
PERIOD(dt, dt2)	WeeksBetween(dt, dt2), YearsBetween (dt, dt2), SecondsBetween (dt, dt2), (dt, dt2), MonthsBetween (dt, dt2), MinutesBetween (dt, dt2) HoursBetween
CAPITALIZE(chaîne)	UCFIRST(chaîne)

Fonctions définies par l'utilisateur abandonnées	Fonction équivalente
CONCAT_WITH()	CONCAT()
INDEX_OF()	INDEXOF()
LAST_INDEX_OF()	LAST_INDEXOF()
SPLIT_ON_REGEX()	STRSPLT()
UNCAPITALIZE()	LCFIRST()

Les éléments suivants UDFs ont été supprimés sans équivalent : FORMAT (), LOCAL_DATE (), LOCAL_TIME (), CENTER (), LEFT_PAD (), REPEAT (), REPLACE_ONCE (), RIGHT_PAD (), STRIP_END (), STRIP_START (), SWAP_CASE ().

Commandes Grunt abandonnées

Certaines commandes Grunt ont été suspendues à commencer par Pig 0.16.0. Le tableau suivant répertorie les commandes Grunt sous Pig 0.14.0 et les commandes équivalentes dans la version actuelle, le cas échéant.

Commandes Grunt Pig 0.14.0 et commandes équivalentes de la version de Pig actuelle

Commande Grunt Pig 0.14.0	Commande Grunt sous Pig 0.16.0 et version ultérieure
chat (< non-hdfs-path >)	fs -cat < non-hdfs-path > ;
cd < non-hdfs-path > ;	Pas d'équivalent
ls < non-hdfs-path > ;	fs -ls < non-hdfs-path > ;
déplacer < non-hdfs-path > < non-hdfs-path > ;	fs -mv < non-hdfs-path > < > ; non-hdfs-path
copier < non-hdfs-path > < non-hdfs-path > ;	fs -cp < non-hdfs-path > < non-hdfs-path > ;
copyToLocal < non-hdfs-path > <local-path>;	fs - copyToLocal < non-hdfs-path > <local-path>;

Commande Grunt Pig 0.14.0	Commande Grunt sous Pig 0.16.0 et version ultérieure
copyFromLocal <local-path>< non-hdfs-path > ;	fs - copyFromLocal <local-path>< non-hdfs-path > ;
mkdir < non-hdfs-path > ;	fs -mkdir < > ; non-hdfs-path
rm < non-hdfs-path > ;	fs -rm -r -SipTrash < > ; non-hdfs-path
rmf < non-hdfs-path > ;	fs -rm -r -SipTrash < > ; non-hdfs-path

Capacités supprimées pour les répertoires de base non HDFS

Pig 0.14.0 sur les versions 4.x d'Amazon EMR possède deux mécanismes pour autoriser les utilisateurs autres que l'utilisateur `hadoop`, qui ne disposent pas de répertoires de base, à exécuter des scripts Pig. Le premier mécanisme est un mécanisme de secours automatique qui définit le répertoire de travail initial dans le répertoire racine, si le répertoire de base n'existe pas. Le second mécanisme est une propriété `pig.initial.fs.name` qui vous permet de modifier le répertoire de travail initial.

Ces mécanismes ne sont pas disponibles à partir des versions Amazon EMR 5.0.0. C'est pourquoi les utilisateurs doivent disposer d'un répertoire de base sur un HDFS. Cela ne s'applique pas à l'utilisateur `hadoop`, car un répertoire de base est provisionné lors du lancement. Les scripts qui s'exécutent à l'aide des étapes JAR Hadoop utilisent par défaut l'utilisateur Hadoop, sauf si un autre utilisateur est explicitement spécifié à l'aide de `command-runner.jar`.

Amazon EMR version 4.9.6

Versions d'application 4.9.6

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notes de mise à jour 4.9.6

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature Version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature Version 4 exclusivement pour authentifier les demandes à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 4.9.6

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.17.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-9	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-9	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	1.0.0-amzn-9	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-9	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-9	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.3	Clients de ligne de commande Spark.
spark-history-server	1.6.3	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.3	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.3	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client de ligne de commande.

Classifications des configurations 4.9.6

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.9.6

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.

Classifications	Description
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4 j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.

Classifications	Description
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.

Classifications	Description
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.

Classifications	Description
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.9.5

Versions d'application 4.9.5

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notes de mise à jour 4.9.5

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.9.5. Les modifications ont été apportées à la version 4.9.4.

Date de parution initiale : 29 août 2018

Modifications, améliorations et problèmes résolus

- HBase
  - Cette version corrige une vulnérabilité de sécurité potentielle.

Versions des composants 4.9.5

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.17.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst

Composant	Version	Description
		ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-9	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-9	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-9	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-9	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-9	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.

Composant	Version	Description
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.3	Clients de ligne de commande Spark.
spark-history-server	1.6.3	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	1.6.3	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.3	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client de ligne de commande.

Classifications des configurations 4.9.5

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.9.5

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.

Classifications	Description
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.

Classifications	Description
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.

Classifications	Description
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.

Classifications	Description
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.9.4

Versions d'application 4.9.4

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.9.4	emr-4.9.3	emr-4.9.2	emr-4.9.1
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.9.4	emr-4.9.3	emr-4.9.2	emr-4.9.1
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notes de mise à jour 4.9.4

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.9.4. Les modifications ont été apportées à la version 4.9.3.

Date de parution initiale : 29 mars 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités potentielles.

Versions des composants 4.9.4

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.17.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.2.2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-9	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-9	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-9	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-9	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.

Composant	Version	Description
hive-server	1.0.0-amzn-9	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.3	Clients de ligne de commande Spark.
spark-history-server	1.6.3	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.3	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.3	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.9	ZooKeeper client de ligne de commande.

Classifications des configurations 4.9.4

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.9.4

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.

Classifications	Description
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.

Classifications	Description
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra .properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.prop erties de Presto.

Classifications	Description
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.

Classifications	Description
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.9.3

Versions d'application 4.9.3

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-

	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-

	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notes de mise à jour 4.9.3

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.9.3. Il s'agit des modifications apportées à la version Amazon EMR 4.9.2.

Date de parution initiale : 22 janvier 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités associées à l'exécution spéculative (CVE-2017-5715, CVE-2017-5753 et CVE-2017-5754). Pour de plus amples informations, veuillez consulter https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versions des composants 4.9.3

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-`

`component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.17.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-9	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-9	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	1.0.0-amzn-9	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-9	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-9	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.3	Clients de ligne de commande Spark.
spark-history-server	1.6.3	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.3	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.3	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client de ligne de commande.

Classifications des configurations 4.9.3

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.9.3

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.

Classifications	Description
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4 j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.

Classifications	Description
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.

Classifications	Description
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.

Classifications	Description
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.9.2

Versions d'application 4.9.2

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notes de mise à jour 4.9.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.9.2. Il s'agit des modifications apportées à la version Amazon EMR 4.9.1.

Date de parution : 13 juillet 2017

Des modifications mineures, des correctifs de bogues et des améliorations ont été apportées à cette version.

Versions des composants 4.9.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont

propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.3.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.17.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.

Composant	Version	Description
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.

Composant	Version	Description
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-9	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	1.0.0-amzn-9	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-9	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-9	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-9	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-9	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.3	Clients de ligne de commande Spark.
spark-history-server	1.6.3	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.3	Moteur d'exécution en mémoire pour YARN.

Composant	Version	Description
spark-yarn-slave	1.6.3	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'application tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client de ligne de commande.

Classifications des configurations 4.9.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.9.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.

Classifications	Description
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.

Classifications	Description
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.

Classifications	Description
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.

Classifications	Description
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.9.1

Versions d'application 4.9.1

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notes de mise à jour 4.9.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.9.1. Il s'agit des modifications apportées à la version Amazon EMR 4.8.4.

Date de parution : 10 avril 2017

Problèmes connus résolus depuis les versions précédentes

- Rétroportages de HIVE-9976 et HIVE-10106
- Résolution d'un problème dans YARN où un nombre important de nœuds (plus de 2 000) et de conteneurs (plus de 5 000) provoquait une out-of-memory erreur, par exemple :`"Exception in thread main java.lang.OutOfMemoryError"`.

Modifications et améliorations

- Les versions d'Amazon EMR sont désormais basées sur Amazon Linux 2017.03. Pour de plus amples informations, veuillez consulter https://aws.amazon.com/amazon-linux-ami/2017.03-release-notes/.
- Suppression de Python 2.6 de l'image Linux de base d'Amazon EMR. Vous pouvez installer Python 2.6 manuellement si nécessaire.

Versions des composants 4.9.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.2.0	Connecteur Amazon DynamoDB pour les applicati ons de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.3.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.15.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.7.3-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.3-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-9	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-9	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	1.0.0-amzn-9	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-9	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-9	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.3	Clients de ligne de commande Spark.
spark-history-server	1.6.3	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.3	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.3	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client de ligne de commande.

Classifications des configurations 4.9.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.9.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.

Classifications	Description
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4 j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.

Classifications	Description
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.

Classifications	Description
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.

Classifications	Description
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.8.5

Versions d'application 4.8.5

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.152.3
Spark	1.6.3	1.6.3	1.6.3	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.8

Notes de mise à jour 4.8.5

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature Version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature Version 4 exclusivement pour authentifier les demandes à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 4.8.5

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.14.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst

Composant	Version	Description
		ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-8	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-8	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-8	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-8	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-8	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-8	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.

Composant	Version	Description
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.3	Clients de ligne de commande Spark.
spark-history-server	1.6.3	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	1.6.3	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.3	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client de ligne de commande.

Classifications des configurations 4.8.5

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.8.5

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.

Classifications	Description
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.

Classifications	Description
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.

Classifications	Description
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.

Classifications	Description
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.8.4

Versions d'application 4.8.4

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.152.3	0.151
Spark	1.6.3	1.6.3	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.8	3.4.8

Notes de mise à jour 4.8.4

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.8.4. Il s'agit des modifications apportées à la version Amazon EMR 4.8.3.

Date de version : 7 février 2017

Des modifications mineures, des correctifs de bogues et des améliorations ont été apportées à cette version.

Versions des composants 4.8.4

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.14.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.

Composant	Version	Description
hadoop-httpfs-server	2.7.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-8	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-8	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-8	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-8	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-8	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-8	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.54+	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.3	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	1.6.3	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.3	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.3	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.25+	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client de ligne de commande.

Classifications des configurations 4.8.4

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.8.4

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.

Classifications	Description
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.

Classifications	Description
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.

Classifications	Description
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile .properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.p roperties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.pro perties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresq l.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.pr operties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.pro perties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.prop erties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-def aults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.p roperties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.

Classifications	Description
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.8.3

Versions d'application 4.8.3

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.152.3	0.151	0.151
Spark	1.6.3	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.8	3.4.8	3.4.8

Notes de mise à jour 4.8.3

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.8.3. Il s'agit des modifications apportées à la version Amazon EMR 4.8.2.

Date de parution : 29 décembre 2016

Mises à niveau

- Mise à niveau vers Presto 0.157.1. Pour plus d'informations, consultez Notes de mise à jour Presto dans la documentation Presto.
- Mise à niveau vers Spark 1.6.3. Pour plus d'informations, consultez les Notes de mise à jour Spark dans la documentation Apache Spark.
- Mis à jour vers la version ZooKeeper 3.4.9. Pour plus d'informations, consultez les ZooKeeperNotes de mise à jour dans la ZooKeeper documentation Apache.

Modifications et améliorations

- Ajout de la prise en charge du type d'instance EC2 m4.16xlarge Amazon EMR, sauf les versions 5.0.0, 5.0.3 et 5.2.0.
- Les versions d'Amazon EMR sont désormais basées sur Amazon Linux 2016.09. Pour de plus amples informations, veuillez consulter https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Problèmes connus résolus depuis les versions précédentes

- Correction d'un problème dans Hadoop où le ReplicationMonitor thread pouvait rester bloqué pendant longtemps en raison d'une course entre la réplication et la suppression du même fichier dans un grand cluster.

- Correction d'un problème en raison duquel ControlledJob #toString échouait avec une exception de pointeur nul (NPE) lorsque le statut de la tâche n'était pas correctement mis à jour.

Versions des composants 4.8.3

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.2.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.

Composant	Version	Description
emrfs	2.13.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.3-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-1	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.7.3-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.

Composant	Version	Description
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-8	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-8	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-8	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-8	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-8	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-8	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop

Composant	Version	Description
mahout-client	0.12.2	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5.52	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.157.1	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.157.1	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.3	Clients de ligne de commande Spark.

Composant	Version	Description
spark-history-server	1.6.3	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.3	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.3	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.9	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.9	ZooKeeper client de ligne de commande.

Classifications des configurations 4.8.3

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.8.3

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.

Classifications	Description
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive

Classifications	Description
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.

Classifications	Description
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.

Classifications	Description
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile .properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.p roperties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.pro perties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresq l.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.pr operties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.pro perties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.prop erties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-def aults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.p roperties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.

Classifications	Description
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.8.2

Versions d'application 4.8.2

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.152.3	0.151	0.151	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0,5.6

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 4.8.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.8.2. Il s'agit des modifications apportées à la version Amazon EMR 4.8.0.

Date de parution : 24 octobre 2016

Mises à niveau

- Mise à niveau vers Hadoop 2.7.3
- Mise à niveau vers Presto 0.152.3, qui comprend la prise en charge de l'interface Web Presto. Utilisez le port 8889 du coordinateur Presto pour accéder à l'interface Web Presto. Pour plus d'informations sur l'interface Web Presto, consultez Interface Web dans la documentation Presto.
- Les versions d'Amazon EMR sont désormais basées sur Amazon Linux 2016.09. Pour de plus amples informations, veuillez consulter https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Versions des composants 4.8.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-`

`component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	4.1.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.10.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.7.3-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.3-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.3-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.3-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.3-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.3-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.3-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.3-amzn-0	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-7	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-7	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-7	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	1.0.0-amzn-7	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-7	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-7	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.52	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.152.3	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.152.3	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.2	Clients de ligne de commande Spark.
spark-history-server	1.6.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client de ligne de commande.

Classifications des configurations 4.8.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.8.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.

Classifications	Description
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4 j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.

Classifications	Description
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.

Classifications	Description
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.

Classifications	Description
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.8.1

Versions d'application 4.8.1

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
Presto-Sandbox	0.151	0.151	0.148	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0,5.6	0,5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 4.8.1

Versions des composants 4.8.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.9.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par

Composant	Version	Description
		le collecteur de métadonnées Ganglia.
hadoop-client	2.7.2-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.2-amzn-4	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.2-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.2-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.2-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-7	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-7	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-7	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	1.0.0-amzn-7	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-7	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-7	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.51	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.151	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.151	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.2	Clients de ligne de commande Spark.
spark-history-server	1.6.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client de ligne de commande.

Classifications des configurations 4.8.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.8.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.

Classifications	Description
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4 j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.

Classifications	Description
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.

Classifications	Description
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole.properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.

Classifications	Description
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.

Classifications	Description
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.8.0

Versions d'application 4.8.0

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
Presto-Sandbox	0.151	0.148	0.148	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0,5.6	0,5.6	0,5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 4.8.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.8.0. Il s'agit des modifications apportées à la version Amazon EMR 4.7.2.

Date de parution : 7 septembre 2016

Mises à niveau

- Mis à jour vers la version HBase 1.2.2
- Mise à niveau vers Presto-Sandbox 0.151
- Mise à niveau vers Tez 0.8.4
- Mise à niveau vers Zeppelin-Sandbox 0.6.1

Modifications et améliorations

- Correction d'un problème dans YARN à cause duquel ils ApplicationMaster tentaient de nettoyer des conteneurs qui n'existent plus parce que leurs instances avaient été fermées.
- Correction de l'URL hive-server2 pour les actions Hive2 dans les exemples Oozie.
- Ajout de la prise en charge d'autres catalogues Presto.
- Correctifs rétroportés : HIVE-8948, HIVE-12679, HIVE-13405, PHOENIX-3116, HADOOP-12689
- Ajout de la prise en charge des configurations de sécurité, qui vous permettent de créer et d'appliquer des options de chiffrement plus facilement. Pour en savoir plus, consultez Chiffrement des données.

Versions des composants 4.8.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.

Composant	Version	Description
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.9.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.2-amzn-4	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-4	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.2-amzn-4	Bibliothèque et client de ligne de commande HDFS

Composant	Version	Description
hadoop-hdfs-namenode	2.7.2-amzn-4	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-4	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-4	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-4	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-4	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.2-amzn-4	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.2-amzn-4	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.2	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.2.2	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.2	HBase client en ligne de commande.
hbase-rest-server	1.2.2	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.2	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-7	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-7	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-7	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-7	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-7	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.

Composant	Version	Description
hive-server	1.0.0-amzn-7	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.51	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.151	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.151	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.2	Clients de ligne de commande Spark.
spark-history-server	1.6.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.4	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.6.1	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.8	ZooKeeper client de ligne de commande.

Classifications des configurations 4.8.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.8.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.

Classifications	Description
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.

Classifications	Description
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hiveserver2-site	Modifiez les valeurs dans le fichier hiveserver2-site.xml de Hive Server2
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.

Classifications	Description
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-connector-blackhole	Modifiez les valeurs dans le fichier blackhole .properties de Presto.
presto-connector-cassandra	Modifiez les valeurs dans le fichier cassandra .properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.prop erties de Presto.

Classifications	Description
presto-connector-jmx	Modifiez les valeurs dans le fichier jmx.properties de Presto.
presto-connector-kafka	Modifiez les valeurs dans le fichier kafka.properties de Presto.
presto-connector-localfile	Modifiez les valeurs dans le fichier localfile.properties de Presto.
presto-connector-mongodb	Modifiez les valeurs dans le fichier mongodb.properties de Presto.
presto-connector-mysql	Modifiez les valeurs dans le fichier mysql.properties de Presto.
presto-connector-postgresql	Modifiez les valeurs dans le fichier postgresql.properties de Presto.
presto-connector-raptor	Modifiez les valeurs dans le fichier raptor.properties de Presto.
presto-connector-redis	Modifiez les valeurs dans le fichier redis.properties de Presto.
presto-connector-tpch	Modifiez les valeurs dans le fichier tpch.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.

Classifications	Description
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.7.4

Versions d'application 4.7.4

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.148	0.148	0.147
Spark	1.6.2	1.6.2	1.6.2	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5.6	0,5.6	0,5.6	0,5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 4.7.4

Il s'agit d'une version de correctif visant à ajouter l'authentification AWS Signature Version 4 pour les demandes adressées à Amazon S3. Toutes les applications et tous les composants sont identiques à ceux de la version précédente d'Amazon EMR.

> **Important**
>
> Dans cette version, Amazon EMR utilise AWS Signature Version 4 exclusivement pour authentifier les demandes à Amazon S3. Pour de plus amples informations, veuillez consulter Nouveautés.

Versions des composants 4.7.4

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-`

`component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.8.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.7.2-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.2-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.2-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hadoop-yarn-timeline-server	2.7.2-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.1	HBase client en ligne de commande.
hbase-rest-server	1.2.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-6	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-6	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-6	Point de terminaison HTTP fournissant une interface REST pour HCatalog.

Composant	Version	Description
hive-client	1.0.0-amzn-6	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-6	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-6	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.46	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client

Composant	Version	Description
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.148	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.148	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.2	Clients de ligne de commande Spark.
spark-history-server	1.6.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.3	Les bibliothèques et l'applica tion tez YARN.

Composant	Version	Description
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.5.6-incubating	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client de ligne de commande.

Classifications des configurations 4.7.4

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.7.4

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.

Classifications	Description
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.

Classifications	Description
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4 j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.

Classifications	Description
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Classifications	Description
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.7.3

Versions d'application 4.7.3

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,75
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.0	0.12.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.148	0.147	0.147
Spark	1.6.2	1.6.2	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5.6	0,5.6	0,5.6	0,5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 4.7.3

Versions des composants 4.7.3

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous

nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.8.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.2-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.2-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.7.2-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.2-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.1	HBase client en ligne de commande.
hbase-rest-server	1.2.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-6	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	1.0.0-amzn-6	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-6	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-6	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-6	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-6	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.46	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.148	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.148	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.2	Clients de ligne de commande Spark.
spark-history-server	1.6.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.2	Moteur d'exécution en mémoire pour YARN.

Composant	Version	Description
spark-yarn-slave	1.6.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.3	Les bibliothèques et l'application tez YARN.
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.5.6-incubating	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client de ligne de commande.

Classifications des configurations 4.7.3

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.7.3

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.

Classifications	Description
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.

Classifications	Description
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.

Classifications	Description
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.7.2

Versions d'application 4.7.2

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,75	1,1,27
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.0
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
Mahout	0.12.2	0.12.0	0.12.0	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.147	0.147	0.143
Spark	1.6.2	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5.6	0,5.6	0,5.6	0,5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 4.7.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.7.2.

Date de parution : 15 juillet 2016

Fonctionnalités

- Mise à niveau vers Mahout 0.12.2
- Mise à niveau vers Presto 0.148
- Mise à niveau vers Spark 1.6.2
- Vous pouvez désormais créer un AWSCredentials fournisseur à utiliser avec EMRFS en utilisant un URI comme paramètre. Pour plus d'informations, voir Création d'un AWSCredentials fournisseur pour EMRFS.
- EMRFS permet maintenant aux utilisateurs de configurer un point de terminaison DynamoDB personnalisé pour les métadonnées de leur vue cohérente à l'aide de la propriété `fs.s3.consistent.dynamodb.endpoint` dans le fichier `emrfs-site.xml`.
- Ajout d'un script dans `/usr/bin` appelé `spark-example`, qui encapsule `/usr/lib/spark/spark/bin/run-example` pour vous permettre d'exécuter des exemples directement. Par exemple, pour exécuter l' SparkPi exemple fourni avec la distribution Spark, vous pouvez l'exécuter `spark-example SparkPi 100` à partir de la ligne de commande ou en `command-runner.jar` tant qu'étape dans l'API.

Problèmes connus résolus depuis les versions précédentes

- Résolution d'un problème où `spark-assembly.jar` pour Oozie n'était pas à l'emplacement approprié quand Spark était également installé, ce qui provoquait un échec du lancement d'applications Spark avec Oozie.
- Résolution d'un problème lié à la journalisation basée sur Spark Log4j dans des conteneurs YARN.

Versions des composants 4.7.2

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous

nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.2.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.1.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.8.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.

Composant	Version	Description
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.2-amzn-3	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-3	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.2-amzn-3	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-3	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-3	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-3	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resour cemanager	2.7.2-amzn-3	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.2-amzn-3	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.1	HBase client en ligne de commande.
hbase-rest-server	1.2.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-6	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	1.0.0-amzn-6	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-6	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-6	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-6	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-6	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.2	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.46	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.148	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.148	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.2	Clients de ligne de commande Spark.
spark-history-server	1.6.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.2	Moteur d'exécution en mémoire pour YARN.

Composant	Version	Description
spark-yarn-slave	1.6.2	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.3	Les bibliothèques et l'application tez YARN.
webserver	2.4.23	Serveur HTTP Apache.
zeppelin-server	0.5.6-incubating	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client de ligne de commande.

Classifications des configurations 4.7.2

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.7.2

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hadoop-ssl-server	Modifier la configuration du serveur ssl hadoop
hadoop-ssl-client	Modifier la configuration du client ssl hadoop
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.

Classifications	Description
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.

Classifications	Description
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.

Classifications	Description
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.7.1

Versions d'application 4.7.1

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
AWS Kit SDK pour Java	1,1,75	1,1,75	1,1,27	1,1,27
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.0	1.2.0
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
Mahout	0.12.0	0.12.0	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.147	0.147	0.143	0.143
Spark	1.6.1	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5.6	0,5.6	0,5.6	0,5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notes de mise à jour 4.7.1

Les notes de mise à jour suivantes contiennent des informations sur Amazon EMR 4.7.1.

Date de parution : 10 juin 2016

Problèmes connus résolus depuis les versions précédentes

- Résolution d'un problème qui augmentait la durée de démarrage des clusters lancés dans un VPC avec des sous-réseaux privés. Ce bogue affectait uniquement les clusters lancés avec la version 4.7.0 d'Amazon EMR.
- Résolution d'un problème où des listes de fichiers dans Amazon EMR n'étaient pas traitées correctement pour les clusters lancés avec la version 4.7.0 d'Amazon EMR.

Versions des composants 4.7.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.1.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-goodies	2.0.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.7.1	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.2-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-2	HDFS node-level service for storing blocks.

Composant	Version	Description
hadoop-hdfs-library	2.7.2-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.2-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.2-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.2-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.

Composant	Version	Description
hbase-region-server	1.2.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.1	HBase client en ligne de commande.
hbase-rest-server	1.2.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-5	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-5	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-5	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-5	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-5	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.

Composant	Version	Description
hive-server	1.0.0-amzn-5	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.12.0	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5.46	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.147	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.147	Service pour exécuter les éléments d'une requête.

Composant	Version	Description
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.1	Clients de ligne de commande Spark.
spark-history-server	1.6.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.3	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.18	Serveur HTTP Apache.
zeppelin-server	0.5.6-incubating	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.

Composant	Version	Description
zookeeper-client	3.4.8	ZooKeeper client de ligne de commande.

Classifications des configurations 4.7.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.7.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.

Classifications	Description
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.

Classifications	Description
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.

Classifications	Description
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-site.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.properties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.

Classifications	Description
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.7.0

Versions d'application 4.7.0

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Tez, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
AWS Kit SDK pour Java	1,1,75	1,1,27	1,1,27	1,1,27
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.0	1.2.0	-
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.0	0.11.1	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.147	0.143	0.143	0.140
Spark	1.6.1	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	-	-	-
Trino (PrestoSQL)	-	-	-	-

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5.6	0,5.6	0,5.6	0,5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	-

Notes de mise à jour 4.7.0

> **Important**
>
> Amazon EMR 4.7.0 est obsolète. Veuillez plutôt utiliser Amazon EMR 4.7.1 ou une version ultérieure.

Date de parution : 2 juin 2016

Fonctionnalités

- Ajout d'Apache Phoenix 4.7.0
- Ajout d'Apache Tez 0.8.3
- Mis à jour vers la version HBase 1.2.1
- Mise à niveau vers Mahout 0.12.0
- Mise à niveau vers Presto 0.147
- Mise à niveau de la AWS SDK pour Java version 1.10.75
- L'indicateur final a été supprimé de la propriété `mapreduce.cluster.local.dir` dans `mapred-site.xml` pour permettre aux utilisateurs d'exécuter Pig en mode local.
- Pilotes JDBC Amazon Redshift disponibles sur le cluster

  Les pilotes JDBC Amazon Redshift sont maintenant inclus dans `/usr/share/aws/redshift/jdbc`. `/usr/share/aws/redshift/jdbc/RedshiftJDBC41.jar` est le pilote Amazon Redshift compatible avec JDBC 4.1 et `/usr/share/aws/redshift/jdbc/`

`RedshiftJDBC4.jar` est le pilote Amazon Redshift compatible avec JDBC 4.0. Pour plus d'informations, consultez la section Configuration d'une connexion JDBC dans le Guide de gestion Amazon Redshift.

- Java 8

  Sauf pour Presto, OpenJDK 1.7 est le JDK par défaut utilisé pour toutes les applications. Cependant, OpenJDK 1.7 et 1.8 sont installés. Pour plus d'informations sur la configuration de `JAVA_HOME` pour les applications, consultez Configuration d'applications pour utiliser Java 8.

Problèmes connus résolus depuis les versions précédentes

- Résolution d'un problème de noyau qui affectait de manière significative les performances sur les volumes HDD à débit optimisé (ST1) EBS pour Amazon EMR dans emr-4.6.0.
- Résolution d'un problème où un cluster échouait si une zone de chiffrement HDFS était spécifiée sans que Hadoop ait été choisi comme application.
- Remplacement de la stratégie d'écriture HDFS par défaut `RoundRobin` par `AvailableSpaceVolumeChoosingPolicy`. Certains volumes n'ont pas été utilisés correctement avec la RoundRobin configuration, ce qui a entraîné la défaillance de nœuds principaux et un HDFS peu fiable.
- Résolution d'un problème lié à l'interface de ligne de commande EMRFS, qui entraînait une exception lors de la création de la table de métadonnées DynamoDB par défaut pour des vues cohérentes.
- Résolution d'un problème de blocage dans EMRFS qui pouvait potentiellement se produire lors d'opérations de changement de nom et de copie en plusieurs parties.
- Correction d'un problème lié à EMRFS qui faisait en sorte que la CopyPart taille par défaut était de 5 Mo. La valeur par défaut est maintenant correctement définie sur 128 Mo.
- Résolution d'un problème lié à la configuration upstart dans Zeppelin qui pouvait vous empêcher potentiellement d'arrêter le service.
- Résolution d'un problème lié à Spark et Zeppelin qui vous empêchait d'utiliser le schéma d'URI `s3a://`, car `/usr/lib/hadoop/hadoop-aws.jar` n'était pas chargé correctement dans leur chemin de classe respectif.
- Rétroportage de HUE-2484.
- J'ai rétroporté un commit depuis Hue 3.9.0 (il n'existe pas de JIRA) pour corriger un problème avec l' HBase exemple de navigateur.
- Rétroportage de HIVE-9073.

Versions des composants 4.7.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.1.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.0.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.2.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.4.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.7.1	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst

Composant	Version	Description
		ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.2-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.2-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.2-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.7.2-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.2-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hadoop-yarn-timeline-server	2.7.2-amzn-2	Service de récupération d'informations actuelles et historiques pour les applicati ons YARN.
hbase-hmaster	1.2.1	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.1	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.1	HBase client en ligne de commande.
hbase-rest-server	1.2.1	Service fournissant un point de terminaison RESTful HTTP pour HBase.

Composant	Version	Description
hbase-thrift-server	1.2.1	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-5	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-5	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-5	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-5	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-5	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-5	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-7	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.12.0	Bibliothèque pour la technolog ie Machine Learning.

Composant	Version	Description
mysql-server	5.5.46	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
phoenix-library	4.7.0- 1.2 HBase	Les bibliothèques phoenix pour serveur et client
phoenix-query-server	4.7.0- 1.2 HBase	Un serveur léger fournissant un accès JDBC, ainsi qu'un accès au format JSON et de tampons de protocole, à l'API Avatica
presto-coordinator	0.147	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.147	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.1	Clients de ligne de commande Spark.
spark-history-server	1.6.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.

Composant	Version	Description
spark-on-yarn	1.6.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
tez-on-yarn	0.8.3	Les bibliothèques et l'applica tion tez YARN.
webserver	2.4.18	Serveur HTTP Apache.
zeppelin-server	0.5.6-incubating	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client de ligne de commande.

Classifications des configurations 4.7.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.7.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.

Classifications	Description
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.

Classifications	Description
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
phoenix-hbase-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Modifiez les valeurs dans le fichier hbase-sit e.xml de Phoenix.
phoenix-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Phoenix.
phoenix-metrics	Modifiez les valeurs dans le fichier hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
tez-site	Modifiez les valeurs dans le fichier tez-site.xml de Tez.

Classifications	Description
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.6.1

Versions d'application 4.6.1

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalogHadoop, Hive, Hue, Mahout, Oozie-Sandbox, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
AWS Kit SDK pour Java	1,1,27	1,1,27	1,1,27	1,1,27
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.0	1.2.0	-	-
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.1
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
Mahout	0.11.1	0.11.1	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.143	0.143	0.140	0.136
Spark	1.6.1	1.6.1	1.6.1	1.6.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5.6	0,5.6	0,5.6	0,5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	-	-

Notes de mise à jour 4.6.1

Versions des composants 4.6.1

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.0.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.0.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.1.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.3.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.6.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.2-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.2-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.2-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider

Composant	Version	Description
hadoop-mapred	2.7.2-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.2-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hbase-hmaster	1.2.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.0	HBase client en ligne de commande.
hbase-rest-server	1.2.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-4	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.

Composant	Version	Description
hcatalog-server	1.0.0-amzn-4	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-4	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-4	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-4	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-4	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-6	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.11.1	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.

Composant	Version	Description
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
presto-coordinator	0.143	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.143	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.1	Clients de ligne de commande Spark.
spark-history-server	1.6.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
webserver	2,4	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.5.6-incubating	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informations de configuration, l'affectation de noms, la synchronisation distribuée et la fourniture de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client de ligne de commande.

Classifications des configurations 4.6.1

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.6.1

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.

Classifications	Description
hbase-env	Changez les valeurs dans HBase l'environ nement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environ nement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hiv e-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.

Classifications	Description
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4 j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.

Classifications	Description
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.prop erties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-def aults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.p roperties de Spark.

Classifications	Description
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.6.0

Versions d'application 4.6.0

Cette version inclut les applications suivantes : Ganglia, HBase, HCatalogHadoop, Hive, Hue, Mahout, Oozie-Sandbox, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, Zeppelin-Sandbox, et ZooKeeper-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR

- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
AWS Kit SDK pour Java	1,1,27	1,1,27	1,1,27	1,1,27
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.0	-	-	-
HCatalog	1.0.0	1.0.0	1.0.0	-
Hadoop	2.7.2	2.7.2	2.7.1	2.7.1
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.1	0.11.1	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.143	0.140	0.136	0.130
Spark	1.6.1	1.6.1	1.6.0	1.6.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5.6	0,5.6	0,5.6	0,5,5
ZooKeeper	-	-	-	-

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
ZooKeeper-Sandbox	3.4.8	-	-	-

Notes de mise à jour 4.6.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.6.0.

- Ajout de HBase 1.2.0
- Ajout de Zookeeper-Sandbox 3.4.8
- Mise à niveau vers Presto-Sandbox 0.143
- Les versions Amazon EMR sont désormais basées sur Amazon Linux 2016.03.0. Pour de plus amples informations, veuillez consulter https://aws.amazon.com/amazon-linux-ami/2016.03-release-notes/.
- Problème affectant les types de volume HDD à débit optimisé (ST1) EBS

  Un problème dans les versions 4.2 et précédentes de noyau Linux affecte de manière significative les performances sur les volumes HDD à débit optimisé (ST1) EBS pour EMR. Cette version (emr-4.6.0) utilise la version de noyau 4.4.5 et est donc concernée. Par conséquent, nous vous recommandons de ne pas utiliser emr-4.6.0 si vous voulez utiliser des volumes EBS st1. Vous pouvez utiliser emr-4.5.0 ou des versions précédentes d'Amazon EMR avec ST1 sans incidence. En outre, nous fournirons le correctif avec les versions futures.
- Version de Python par défaut

  Python 3.4 est maintenant installé par défaut, mais Python 2.7 reste la valeur système par défaut. Vous pouvez configurer Python 3.4 comme valeur par défaut du système à l'aide d'une action bootstrap ; vous pouvez utiliser l'API de configuration pour définir l'exportation de PYSPARK_PYTHON sur `/usr/bin/python3.4` dans la classification afin d'affecter la `spark-env` version de Python utilisée par. PySpark
- Java 8

  Sauf pour Presto, OpenJDK 1.7 est le JDK par défaut utilisé pour toutes les applications. Cependant, OpenJDK 1.7 et 1.8 sont installés. Pour plus d'informations sur la configuration de `JAVA_HOME` pour les applications, consultez Configuration d'applications pour utiliser Java 8.

Problèmes connus résolus depuis les versions précédentes

- Résolution d'un problème où le provisionnement d'application échouait parfois de manière aléatoire en raison d'un mot de passe généré.
- Auparavant, `mysqld` était installé sur tous les nœuds. Maintenant, il est uniquement installé sur l'instance principale et seulement si l'application choisie inclut `mysql-server` comme composant. Actuellement, les applications suivantes incluent le `mysql-server` composant : Hive HCatalog, Hue, Presto-Sandbox et Sqoop-Sandbox.
- Remplacement de la valeur par défaut 80 par 32 pour `yarn.scheduler.maximum-allocation-vcores`, ce qui résout un problème introduit dans emr-4.4.0 qui se produit principalement avec Spark lors de l'utilisation de l'option `maximizeResourceAllocation` dans un cluster dont le type d'instance principal est l'un des types d'instance large pour lesquels des cœurs virtuels YARN sont définis sur des valeurs supérieures à 32. Les types c4.8xlarge, cc2.8xlarge, hs1.8xlarge, i2.8xlarge, m2.4xlarge, r3.8xlarge, d2.8xlarge et m4.10xlarge étaient affectés par ce problème.
- s3-dist-cp utilise désormais EMRFS pour toutes les nominations Amazon S3 et n'effectue plus une copie intermédiaire dans un répertoire HDFS temporaire.
- Résolution d'un problème lié au traitement des exceptions pour le chiffrement côté client du chargement partitionné.
- Ajout d'une option pour permettre aux utilisateurs de modifier la classe de stockage Amazon S3. Par défaut, ce paramètre est `STANDARD`. La configuration de classification `emrfs-site` est `fs.s3.storageClass`, et les valeurs possibles sont `STANDARD`, `STANDARD_IA` et `REDUCED_REDUNDANCY`. Pour plus d'informations sur les classes de stockage, consultez la section Classes de stockage dans le Guide de l'utilisateur Amazon Simple Storage Service.

Versions des composants 4.6.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *CommunityVersion*-amzn-*EmrVersion*. *EmrVersion* commence à 0. Par exemple, si un composant de la communauté open source nommé myapp-component avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom 2.2-amzn-2.

Composant	Version	Description
emr-ddb	3.0.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.0.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.1.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.3.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.6.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par

Composant	Version	Description
		le collecteur de métadonnées Ganglia.
hadoop-client	2.7.2-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.2-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.2-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.2-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.

Composant	Version	Description
hbase-hmaster	1.2.0	Service pour un HBase cluster chargé de la coordination des régions et de l'exécution des commandes administratives.
hbase-region-server	1.2.0	Service destiné à desservir une ou plusieurs HBase régions.
hbase-client	1.2.0	HBase client en ligne de commande.
hbase-rest-server	1.2.0	Service fournissant un point de terminaison RESTful HTTP pour HBase.
hbase-thrift-server	1.2.0	Service fournissant un point de terminaison Thrift pour. HBase
hcatalog-client	1.0.0-amzn-4	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-4	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-4	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-4	Client de ligne de commande Hive.

Composant	Version	Description
hive-metastore-server	1.0.0-amzn-4	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-4	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-6	Application web pour l'analyse des données à l'aide d'applica tions de l'écosystème Hadoop
mahout-client	0.11.1	Bibliothèque pour la technolog ie Machine Learning.
mysql-server	5.5	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
presto-coordinator	0.143	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.143	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.

Composant	Version	Description
spark-client	1.6.1	Clients de ligne de commande Spark.
spark-history-server	1.6.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
webserver	2,4	Serveur HTTP Apache.
zeppelin-server	0.5.6-incubating	Portable basé sur le web qui permet l'analyse de données interactives.
zookeeper-server	3.4.8	Service centralisé conçu pour la conservation des informati ons de configuration, l'affecta tion de noms, la synchroni sation distribuée et la fournitur e de services de groupe.
zookeeper-client	3.4.8	ZooKeeper client de ligne de commande.

Classifications des configurations 4.6.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.6.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hbase-env	Changez les valeurs dans HBase l'environnement.
hbase-log4j	Modifiez les valeurs dans le fichier HBase hbase-log4j.properties.
hbase-metrics	Modifiez les valeurs dans le fichier HBase hadoop-metrics2-hbaase.properties.
hbase-policy	Modifiez les valeurs dans HBase le fichier hbase-policy.xml.
hbase-site	Modifiez les valeurs dans HBase le fichier hbase-site.xml.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.

Classifications	Description
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.

Classifications	Description
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.

Classifications	Description
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Classifications	Description
zookeeper-config	Modifiez les valeurs dans ZooKeeper le fichier zoo.cfg.
zookeeper-log4j	Modifiez les valeurs dans le ZooKeeper fichier log4j.properties.

Amazon EMR version 4.5.0

Versions d'application 4.5.0

Cette version inclut les applications suivantes : Ganglia, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, et Zeppelin-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
AWS Kit SDK pour Java	1,1,27	1,1,27	1,1,27	1,1,27
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie

	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.6.0
HBase	-	-	-	-
HCatalog	1.0.0	1.0.0	-	-
Hadoop	2.7.2	2.7.1	2.7.1	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.1	0.11.0	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-

	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.140	0.136	0.130	0.125
Spark	1.6.1	1.6.0	1.6.0	1.5.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5.6	0,5.6	0,5,5	0,5,5
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notes de mise à jour 4.5.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.5.0.

Date de parution : 4 avril 2016

Fonctionnalités

- Mise à niveau vers Spark 1.6.1
- Mise à niveau vers Hadoop 2.7.2

- Mise à niveau vers Presto 0.140
- Ajout de la AWS KMS prise en charge par du chiffrement côté serveur côté serveur Amazon S3.

Problèmes connus résolus depuis les versions précédentes

- Résolution d'un problème où les serveurs MySQL et Apache ne démarraient pas après le redémarrage d'un nœud.
- Résolution d'un problème où IMPORT ne fonctionnait pas correctement avec les tables non partitionnées stockées dans Amazon S3
- Résolution d'un problème lié à Presto où le répertoire intermédiaire devait être `/mnt/tmp` au lieu de `/tmp` lors de l'écriture dans des tables Hive.

Versions des composants 4.5.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.0.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.

Composant	Version	Description
emr-goodies	2.0.0	Bibliothèques proposant plus de commodités pour l'écosyst ème Hadoop.
emr-kinesis	3.1.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.2.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.5.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.2-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.2-amzn-0	HDFS node-level service for storing blocks.

Composant	Version	Description
hadoop-hdfs-library	2.7.2-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.2-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.2-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.2-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.2-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.7.2-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hcatalog-client	1.0.0-amzn-4	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-4	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.

Composant	Version	Description
hcatalog-webhcat-server	1.0.0-amzn-4	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-4	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-4	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-4	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-5	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.11.1	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.

Composant	Version	Description
presto-coordinator	0.140	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.140	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.1	Clients de ligne de commande Spark.
spark-history-server	1.6.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.1	Bibliothèques Apache Spark requises par les esclaves YARN.
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
webserver	2,4	Serveur HTTP Apache.
zeppelin-server	0.5.6-incubating	Portable basé sur le web qui permet l'analyse de données interactives.

Classifications des configurations 4.5.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.5.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.

Classifications	Description
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4 j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.

Classifications	Description
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.

Classifications	Description
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Amazon EMR version 4.4.0

Versions d'application 4.4.0

Cette version inclut les applications suivantes : Ganglia, HCatalog, HadoopHive, Hue, Mahout, Oozie-Sandbox, Pig, Presto-Sandbox, Spark, Sqoop-Sandbox, et Zeppelin-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR

- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
AWS Kit SDK pour Java	1,1,27	1,1,27	1,1,27	Non suivie
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.6.0	-
HBase	-	-	-	-
HCatalog	1.0.0	-	-	-
Hadoop	2.7.1	2.7.1	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.0	0.11.0	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.0.1
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.136	0.130	0.125	0.119
Spark	1.6.0	1.6.0	1.5.2	1.5.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5.6	0,5,5	0,5,5	0.6.0-SNAPSHOT
ZooKeeper	-	-	-	-

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
ZooKeeper-Sandbox	-	-	-	-

Notes de mise à jour 4.4.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.4.0.

Date de parution : 14 mars 2016

Fonctionnalités

- Ajout de HCatalog 1.0.0
- Ajout de Sqoop-Sandbox 1.4.6
- Mise à niveau vers Presto 0.136
- Mise à niveau vers Zeppelin 0.5.6
- Mise à niveau vers Mahout 0.11.1
- `dynamicResourceAllocation` activé par défaut.
- Ajout d'un tableau de toutes les classifications de configuration pour la version. Pour en savoir plus, consultez le tableau Classifications des configurations dans Configuration des applications.

Problèmes connus résolus depuis les versions précédentes

- Correction d'un problème à cause duquel le `maximizeResourceAllocation` paramètre ne réservait pas suffisamment de mémoire pour les ApplicationMaster démons YARN.
- Résolution d'un problème rencontré avec un DNS personnalisé. Si des entrées dans `resolve.conf` précèdent les entrées personnalisées fournies, les entrées personnalisées ne peuvent pas être résolues. Ce comportement était affecté par des clusters dans un VPC dans lequel le serveur de noms VPC par défaut était inséré comme première entrée dans `resolve.conf`.
- Résolution d'un problème où la version Python par défaut passait à la version 2.7 et boto n'était pas installé pour cette version.
- Résolution d'un problème où des conteneurs YARN et des applications Spark généraient un fichier rdd (round robin database) unique Ganglia, si bien que le premier disque attaché à l'instance se

remplissait. En raison de ce correctif, les métriques de niveau conteneur YARN ont été désactivées et les métriques de niveau application Spark ont été désactivées.

- Résolution d'un problème dans le transmetteur de journaux où celui-ci supprimait tous les dossiers de journal vides. De ce fait, l'interface de ligne de commande Hive ne pouvait pas journaliser, car le transmetteur de journaux supprimait le dossier `user` vide sous `/var/log/hive`.
- Résolution d'un problème affectant les importations Hive qui avait une incidence sur le partitionnement et entraînait une erreur lors de l'importation.
- Résolution d'un problème où EMRFS et s3-dist-cp ne géraient pas correctement les noms de compartiment qui contenaient des points.
- Modification d'un comportement dans EMRFS de sorte que dans les compartiments activés pour la gestion des versions, le fichier marqueur `_$folder$` ne soit pas créé en permanence, ce qui peut contribuer à l'amélioration des performances de tels compartiments.
- Modification d'un comportement dans EMRFS pour ne pas utiliser pas des fichiers d'instruction à l'exception des cas où le chiffrement côté client est activé. Si vous souhaitez supprimer des fichiers d'instruction tout en utilisant le chiffrement côté client, vous pouvez définir la propriété d'emrfs-site.xml `fs.s3.cse.cryptoStorageMode.deleteInstructionFiles.enabled` sur true.
- Modification de l'agrégation de journaux YARN pour conserver les journaux dans la destination d'agrégation pendant deux jours. La destination par défaut est le stockage HDFS de votre cluster. Si vous souhaitez modifier cette durée, remplacez la valeur `yarn.log-aggregation.retain-seconds` à l'aide de la classification de configuration `yarn-site` lorsque vous créez votre cluster. Comme toujours, vous pouvez enregistrer vos journaux d'applications dans Amazon S3 à l'aide du paramètre `log-uri` lorsque vous créez votre cluster.

Correctifs appliqués

- HIVE-9655
- HIVE-9183
- HADOOP-12810

Versions des composants 4.4.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon

EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.0.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.0.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.1.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.2.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.4.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonnées Ganglia pour agréger les

Composant	Version	Description
		métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.1-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.1-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.1-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.1-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.1-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.7.1-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.1-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.1-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.

Composant	Version	Description
hadoop-yarn-resourcemanager	2.7.1-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hcatalog-client	1.0.0-amzn-3	Client de ligne de commande « hcat » pour la manipulation de hcatalog-server.
hcatalog-server	1.0.0-amzn-3	Fourniture de services HCatalog, table et couche de gestion du stockage pour les applications distribuées.
hcatalog-webhcat-server	1.0.0-amzn-3	Point de terminaison HTTP fournissant une interface REST pour HCatalog.
hive-client	1.0.0-amzn-3	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-3	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-3	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-5	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.11.1	Bibliothèque pour la technologie Machine Learning.

Composant	Version	Description
mysql-server	5.5	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
presto-coordinator	0.136	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.136	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.0	Clients de ligne de commande Spark.
spark-history-server	1.6.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.0	Bibliothèques Apache Spark requises par les esclaves YARN.

Composant	Version	Description
sqoop-client	1.4.6	Client de ligne de commande Apache Sqoop.
webserver	2,4	Serveur HTTP Apache.
zeppelin-server	0.5.6-incubating	Portable basé sur le web qui permet l'analyse de données interactives.

Classifications des configurations 4.4.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.4.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Hadoop.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.

Classifications	Description
hcatalog-env	Changez les valeurs dans HCatalog l'environnement.
hcatalog-server-jndi	Modifiez les valeurs dans le fichier HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Modifiez les valeurs dans le HCatalog proto-hive-site fichier .xml.
hcatalog-webhcat-env	Modifiez les valeurs dans HCat l'environnement HCatalog Web.
hcatalog-webhcat-log4j	Modifiez les valeurs dans HCatalog HCat log4j.properties du Web.
hcatalog-webhcat-site	Modifiez les valeurs dans HCatalog le fichier webhcat-site.xml HCat de Web.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.

Classifications	Description
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.

Classifications	Description
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
sqoop-env	Modifiez les valeurs d'environnement Sqoop.
sqoop-oraoop-site	Modifiez les valeurs dans le fichier oraoop-site.xml de OraOop Sqoop.
sqoop-site	Modifiez les valeurs dans le fichier sqoop-site.xml de Sqoop.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Amazon EMR version 4.3.0

Versions d'application 4.3.0

Cette version inclut les applications suivantes : GangliaHadoop, Hive, Hue, Mahout, Oozie-Sandbox, Pig, Presto-Sandbox, Spark, et Zeppelin-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK pour Java	1,1,27	1,1,27	Non suivie	Non suivie
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5,5	0,5,5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notes de mise à jour 4.3.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.3.0.

Date de parution : 19 janvier 2016

Fonctionnalités

- Mise à niveau vers Hadoop 2.7.1
- Mise à niveau vers Spark 1.6.0
- Mise à niveau de Ganglia vers la version 3.7.2
- Mise à niveau vers de Presto vers la version 0.130
- Amazon EMR a apporté quelques modifications au paramètre `spark.dynamicAllocation.enabled` lorsque celui-ci est défini sur true ; la valeur par défaut est false. Lorsque ce paramètre est défini sur true, cela a un effet sur les valeurs par défaut définies par le paramètre `maximizeResourceAllocation`.
  - Si `spark.dynamicAllocation.enabled` est défini sur true, `spark.executor.instances` n'est pas défini par `maximizeResourceAllocation`.

- Le paramètre `spark.driver.memory` est désormais configuré en fonction des types d'instances du cluster d'une manière similaire à la façon dont `spark.executors.memory` est défini. Cependant, étant donné que l'application de pilote Spark peut s'exécuter soit sur le maître, soit sur une des instances principales (par exemple, dans le client YARN et les modes cluster, respectivement), le paramètre `spark.driver.memory` est défini en fonction du type d'instance du plus petit des types d'instances entre ces deux groupes d'instances.
- Le paramètre `spark.default.parallelism` est désormais défini sur deux fois le nombre de cœurs de processeurs disponibles pour les conteneurs YARN. Dans les versions précédentes, c'était la moitié de cette valeur.
- Les calculs de la surcharge de mémoire réservée pour les processus Spark YARN ont été ajustés pour être plus précis, ce qui se traduit par une petite augmentation de la quantité de mémoire disponible pour Spark (c'est-à-dire, `spark.executor.memory`).

Problèmes connus résolus depuis les versions précédentes

- L'agrégation de journaux YARN est maintenant activée par défaut.
- Correction d'un problème où les journaux ne pouvaient pas être transmis vers le compartiment de journaux Amazon S3 pour le cluster lorsque l'agrégation de journaux YARN était activée.
- Les tailles de conteneur YARN ont maintenant un nouveau minimum de 32 sur tous les types de nœuds.
- Résolution d'un problème lié à Ganglia qui entraînait des E/S de disque excessives sur le nœud primaire dans les clusters de grande taille.
- Résolution d'un problème qui empêchait les journaux des applications d'être transmis à Amazon S3 lorsqu'un cluster s'arrêtait.
- Résolution d'un problème dans l'interface de ligne de commande EMRFS qui entraînait l'échec de certaines commandes.
- Correction d'un problème lié à Zeppelin qui empêchait le chargement des dépendances dans le sous-jacent. SparkContext
- Résolution d'un problème provoqué par une tentative de redimensionnement pour ajouter des instances.
- Résolution d'un problème dans Hive où CREATE TABLE AS SELECT effectuait des appels de liste excessifs vers Amazon S3.
- Résolution d'un problème où les grands clusters n'étaient pas provisionnés correctement lorsque Hue, Oozie et Ganglia étaient installés.

- Résolution d'un problème dans s3-dist-cp où un code de sortie zéro était renvoyé même en cas d'échec avec une erreur.

Versions des composants 4.3.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.0.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.0.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.1.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.1.0	Application de copie distribuée optimisée pour Amazon S3.

Composant	Version	Description
emrfs	2.3.0	Connecteur Amazon S3 pour les applications de l'écosyst ème Hadoop.
ganglia-monitor	3.7.2	Agent Ganglia intégré pour les applications de l'écosyst ème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.7.2	Collecteur de métadonné es Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.7.1	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.
hadoop-client	2.7.1-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.7.1-amzn-0	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.7.1-amzn-0	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.7.1-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.7.1-amzn-0	Point de terminaison HTTP pour les opérations HDFS.

Composant	Version	Description
hadoop-kms-server	2.7.1-amzn-0	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.7.1-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.7.1-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resourcemanager	2.7.1-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hive-client	1.0.0-amzn-2	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-2	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-2	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-5	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.11.0	Bibliothèque pour la technologie Machine Learning.

Composant	Version	Description
mysql-server	5.5	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
presto-coordinator	0.130	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.130	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.6.0	Clients de ligne de commande Spark.
spark-history-server	1.6.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.6.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.6.0	Bibliothèques Apache Spark requises par les esclaves YARN.
webserver	2,4	Serveur HTTP Apache.

Composant	Version	Description
zeppelin-server	0.5.5-incubating-amzn-1	Portable basé sur le web qui permet l'analyse de données interactives.

Classifications des configurations 4.3.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.3.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.

Classifications	Description
hive-log4j	Modifiez les valeurs dans le fichier hive-log4j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-site.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls.xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j.properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-site.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log4j.properties d'Oozie.

Classifications	Description
oozie-site	Modifiez les valeurs dans le fichier oozie-site.xml d'Oozie.
pig-properties	Modifiez les valeurs dans le fichier pig.properties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.properties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.properties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.properties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.

Classifications	Description
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Amazon EMR version 4.2.0

Versions d'application 4.2.0

Cette version inclut les applications suivantes : GangliaHadoop, Hive, Hue, Mahout, Oozie-Sandbox, Pig, Presto-Sandbox, Spark, et Zeppelin-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK pour Java	1,1,27	1,1,27	Non suivie	Non suivie
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5,5	0,5,5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notes de mise à jour 4.2.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.2.0.

Date de parution : 18 novembre 2015

Fonctionnalités

- Ajout de la prise en charge de Ganglia
- Mise à niveau vers Spark 1.5.2
- Mise à niveau vers Presto 0.125
- Mise à niveau d'Oozie vers la version 4.2.0
- Mise à niveau de Zeppelin vers la version 0.5.5
- Mise à niveau de la AWS SDK pour Java version 1.10.27

Problèmes connus résolus depuis les versions précédentes

- Résolution d'un problème où l'interface de ligne de commande EMRFS n'utilisait pas le nom de table de métadonnées par défaut.
- Résolution d'un problème rencontré lors de l'utilisation de tables basée sur ORC dans Amazon S3.
- Résolution d'un problème rencontré avec une incompatibilité de version Python dans la configuration de Spark.
- Résolution d'un problème où un état de nœud YARN n'était pas signalé en raison de problèmes de DNS pour des clusters dans un VPC.
- Correction d'un problème rencontré lorsque YARN mettait des nœuds hors service, ce qui entraînait des applications bloquées ou l'incapacité de planifier de nouvelles applications.
- Résolution d'un problème rencontré lorsque des clusters prenaient fin avec l'état TIMED_OUT_STARTING.
- Résolution d'un problème rencontré lors de l'inclusion de la dépendance EMRFS Scala dans d'autres versions. La dépendance Scala a été supprimée.

Versions des composants 4.2.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.0.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.0.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.1.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.0.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.2.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
ganglia-monitor	3.6.0	Agent Ganglia intégré pour les applications de l'écosystème Hadoop avec agent de surveillance Ganglia.
ganglia-metadata-collector	3.6.0	Collecteur de métadonnées Ganglia pour agréger les métriques des agents de surveillance Ganglia.
ganglia-web	3.5.10	Application web pour afficher les métriques collectées par le collecteur de métadonnées Ganglia.

Composant	Version	Description
hadoop-client	2.6.0-amzn-2	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.6.0-amzn-2	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.6.0-amzn-2	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.6.0-amzn-2	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.6.0-amzn-2	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.6.0-amzn-2	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.6.0-amzn-2	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.6.0-amzn-2	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resourcemanager	2.6.0-amzn-2	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hive-client	1.0.0-amzn-1	Client de ligne de commande Hive.

Composant	Version	Description
hive-metastore-server	1.0.0-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-5	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.11.0	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5	Serveur de base de données MySQL.
oozie-client	4.2.0	Client de ligne de commande Oozie.
oozie-server	4.2.0	Service pour l'acceptation des demandes de flux de travail Oozie.
presto-coordinator	0.125	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.125	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.

Composant	Version	Description
spark-client	1.5.2	Clients de ligne de commande Spark.
spark-history-server	1.5.2	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.5.2	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.5.2	Bibliothèques Apache Spark requises par les esclaves YARN.
webserver	2,4	Serveur HTTP Apache.
zeppelin-server	0.5.5-incubating-amzn-0	Portable basé sur le web qui permet l'analyse de données interactives.

Classifications des configurations 4.2.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.2.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.

Classifications	Description
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4 j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.

Classifications	Description
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
presto-connector-hive	Modifiez les valeurs dans le fichier hive.prop erties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.

Classifications	Description
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
spark-metrics	Modifiez les valeurs dans le fichier metrics.properties de Spark.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Amazon EMR version 4.1.0

Versions d'application 4.1.0

Cette version inclut les applications suivantes : HadoopHive, Hue, Mahout, Oozie-Sandbox, Pig, Presto-Sandbox, Spark, et Zeppelin-Sandbox.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR

- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK pour Java	1,1,27	1,1,27	Non suivie	Non suivie
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5,5	0,5,5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notes de mise à jour 4.1.0

Versions des composants 4.1.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.0.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.0.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.1.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.0.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.1.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.

Composant	Version	Description
hadoop-client	2.6.0-amzn-1	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.6.0-amzn-1	HDFS node-level service for storing blocks.
hadoop-hdfs-library	2.6.0-amzn-1	Bibliothèque et client de ligne de commande HDFS
hadoop-hdfs-namenode	2.6.0-amzn-1	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.6.0-amzn-1	Point de terminaison HTTP pour les opérations HDFS.
hadoop-kms-server	2.6.0-amzn-1	Serveur de gestion des clés cryptographiques basé sur l'API de Hadoop. KeyProvider
hadoop-mapred	2.6.0-amzn-1	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.6.0-amzn-1	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resourcemanager	2.6.0-amzn-1	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hive-client	1.0.0-amzn-1	Client de ligne de commande Hive.

Composant	Version	Description
hive-metastore-server	1.0.0-amzn-1	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-1	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
hue-server	3.7.1-amzn-4	Application web pour l'analyse des données à l'aide d'applications de l'écosystème Hadoop
mahout-client	0.11.0	Bibliothèque pour la technologie Machine Learning.
mysql-server	5.5	Serveur de base de données MySQL.
oozie-client	4.0.1	Client de ligne de commande Oozie.
oozie-server	4.0.1	Service pour l'acceptation des demandes de flux de travail Oozie.
presto-coordinator	0.119	Service pour accepter les requêtes et gérer l'exécution des requêtes des composants presto-worker.
presto-worker	0.119	Service pour exécuter les éléments d'une requête.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.

Composant	Version	Description
spark-client	1.5.0	Clients de ligne de commande Spark.
spark-history-server	1.5.0	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.5.0	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.5.0	Bibliothèques Apache Spark requises par les esclaves YARN.
zeppelin-server	0.6.0-incubating-SNAPSHOT	Portable basé sur le web qui permet l'analyse de données interactives.

Classifications des configurations 4.1.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.1.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.
emrfs-site	Modifiez les paramètres EMRFS.

Classifications	Description
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hdfs-encryption-zones	Configurez les zones de chiffrement HDFS.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4 j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
hue-ini	Modifiez les valeurs dans le fichier ini de Hue
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Modifiez les valeurs dans le fichier kms-acls. xml de Hadoop.
hadoop-kms-env	Modifiez les valeurs dans l'environnement KMS de Hadoop.
hadoop-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Hadoop.

Classifications	Description
hadoop-kms-site	Modifiez les valeurs dans le fichier kms-site.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
oozie-env	Modifiez les valeurs dans l'environnement d'Oozie.
oozie-log4j	Modifier les valeurs dans le fichier oozie-log 4j.properties d'Oozie.
oozie-site	Modifiez les valeurs dans le fichier oozie-sit e.xml d'Oozie.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.
pig-log4j	Modifiez les valeurs dans le fichier log4j.pro perties de Pig.
presto-log	Modifiez les valeurs dans le fichier log.prope rties de Presto.
presto-config	Modifiez les valeurs dans le fichier config.pr operties de Presto.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-def aults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.

Classifications	Description
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.
zeppelin-env	Modifiez les valeurs dans l'environnement Zeppelin.

Amazon EMR version 4.0.0

Versions d'application 4.0.0

Cette version inclut les applications suivantes : HadoopHive, Mahout, Pig, et Spark.

Le tableau ci-dessous répertorie les versions d'application disponibles dans cette version d'Amazon EMR et les versions d'application des trois versions précédentes d'Amazon EMR (le cas échéant).

Pour obtenir un historique complet des versions des applications de chaque version d'Amazon EMR, consultez les rubriques suivantes :

- Versions des applications dans les versions 7.x d'Amazon EMR
- Versions des applications dans les versions 6.x d'Amazon EMR
- Versions des applications dans les versions 5.x d'Amazon EMR
- Versions des applications dans les versions 4.x d'Amazon EMR

Informations sur la version de l'application

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK pour Java	1,1,27	1,1,27	Non suivie	Non suivie

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Python	Non suivie	Non suivie	Non suivie	Non suivie
Scala	Non suivie	Non suivie	Non suivie	Non suivie
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0,5,5	0,5,5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notes de mise à jour 4.0.0

Versions des composants 4.0.0

Les composants installés par Amazon EMR avec cette version sont répertoriés ci-dessous. Certains sont installés dans le cadre de packages d'application de Big Data. Les autres sont

propres à Amazon EMR et installés pour les fonctions et processus système. Ceux-ci commencent généralement par `emr` ou `aws`. Les packages d'application de big data de la version Amazon EMR la plus récente sont généralement la dernière version trouvée dans la communauté. Nous nous efforçons de mettre à disposition les versions de la communauté dans Amazon EMR le plus rapidement possible.

Certains composants dans Amazon EMR diffèrent des versions de la communauté. Ces composants ont une étiquette de version sous la forme *`CommunityVersion`*`-amzn-`*`EmrVersion`*. *`EmrVersion`* commence à 0. Par exemple, si un composant de la communauté open source nommé `myapp-component` avec la version 2.2 a été modifié trois fois en vue de son inclusion dans différentes versions d'Amazon EMR, sa version apparaît sous le nom `2.2-amzn-2`.

Composant	Version	Description
emr-ddb	3.0.0	Connecteur Amazon DynamoDB pour les applications de l'écosystème Hadoop.
emr-goodies	2.0.0	Bibliothèques proposant plus de commodités pour l'écosystème Hadoop.
emr-kinesis	3.0.0	Connecteur Amazon Kinesis pour les applications de l'écosystème Hadoop.
emr-s3-dist-cp	2.0.0	Application de copie distribuée optimisée pour Amazon S3.
emrfs	2.0.0	Connecteur Amazon S3 pour les applications de l'écosystème Hadoop.
hadoop-client	2.6.0-amzn-0	Clients de ligne de commande Hadoop tels que « hdfs », « hadoop » ou « yarn ».
hadoop-hdfs-datanode	2.6.0-amzn-0	HDFS node-level service for storing blocks.

Composant	Version	Description
hadoop-hdfs-namenode	2.6.0-amzn-0	Service HDFS pour le suivi des noms de fichier et des emplacements de bloc.
hadoop-httpfs-server	2.6.0-amzn-0	Point de terminaison HTTP pour les opérations HDFS.
hadoop-mapred	2.6.0-amzn-0	MapReduce bibliothèques de moteurs d'exécution pour exécuter une MapReduce application.
hadoop-yarn-nodemanager	2.6.0-amzn-0	Service YARN pour la gestion de conteneurs sur un nœud individuel.
hadoop-yarn-resour cemanager	2.6.0-amzn-0	Service YARN pour l'allocation et la gestion des ressources de cluster et des applications distribuées.
hive-client	1.0.0-amzn-0	Client de ligne de commande Hive.
hive-metastore-server	1.0.0-amzn-0	Service pour accéder au metastore Hive, référentiel sémantique stockant des métadonnées pour SQL sur les opérations Hadoop.
hive-server	1.0.0-amzn-0	Service pour l'acceptation de requêtes Hive en tant que requêtes web.
mahout-client	0.10.0	Bibliothèque pour la technolog ie Machine Learning.

Composant	Version	Description
mysql-server	5.5	Serveur de base de données MySQL.
pig-client	0.14.0-amzn-0	Client de ligne de commande Pig.
spark-client	1.4.1	Clients de ligne de commande Spark.
spark-history-server	1.4.1	Interface web pour afficher des événements enregistrés pour la durée de vie d'une application Spark terminée.
spark-on-yarn	1.4.1	Moteur d'exécution en mémoire pour YARN.
spark-yarn-slave	1.4.1	Bibliothèques Apache Spark requises par les esclaves YARN.

Classifications des configurations 4.0.0

Les classifications de configuration vous permettent de personnaliser les applications. Elles correspondent souvent à un fichier XML de configuration de l'application, tel que `hive-site.xml`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Classifications emr-4.0.0

Classifications	Description
capacity-scheduler	Modifiez les valeurs dans le fichier capacity-scheduler.xml de Hadoop.
core-site	Modifiez les valeurs dans le fichier core-site.xml de Hadoop.

Classifications	Description
emrfs-site	Modifiez les paramètres EMRFS.
hadoop-env	Modifiez les valeurs dans l'environnement Hadoop pour tous les composants Hadoop.
hadoop-log4j	Modifiez les valeurs dans le fichier log4j.properties de Hadoop.
hdfs-site	Modifiez les valeurs dans le fichier hdfs-site.xml de HDFS.
hive-env	Modifiez les valeurs dans l'environnement Hive.
hive-exec-log4j	Modifiez les valeurs dans le fichier hive-exec-log 4j.properties de Hive.
hive-log4j	Modifiez les valeurs dans le fichier hive-log4 j.properties de Hive.
hive-site	Modifiez les valeurs dans le fichier hive-site.xml de Hive
httpfs-env	Modifiez les valeurs dans l'environnement HTTPFS.
httpfs-site	Modifiez les valeurs dans le fichier httpfs-si te.xml de Hadoop.
mapred-env	Modifiez les valeurs dans l'environnement de l' MapReduce application.
mapred-site	Modifiez les valeurs dans le fichier mapred-si te.xml de l' MapReduce application.
pig-properties	Modifiez les valeurs dans le fichier pig.prope rties de Pig.

Classifications	Description
pig-log4j	Modifiez les valeurs dans le fichier log4j.properties de Pig.
spark	Paramètres définis par Amazon EMR pour Apache Spark.
spark-defaults	Modifiez les valeurs dans le fichier spark-defaults.conf de Spark.
spark-env	Modifiez les valeurs dans l'environnement Spark.
spark-log4j	Modifiez les valeurs dans le fichier log4j.properties de Spark.
yarn-env	Modifiez les valeurs dans l'environnement YARN.
yarn-site	Modifiez les valeurs dans le fichier yarn-site.xml de YARN.

Versions AMI 2.x et 3.x d'Amazon EMR

> **Note**
>
> AWS met à jour la configuration TLS pour tous les points de terminaison d' AWS API, avec une version minimale de TLS 1.2. Les versions 3.10 et antérieures d'Amazon EMR ne prennent en charge que les connexions TLS 1.0/1.1. Après le 4 décembre 2023, vous ne pourrez plus créer de clusters avec Amazon EMR 3.10 ou version antérieure.
> Si vous utilisez Amazon EMR 3.10 ou une version antérieure, nous vous recommandons de tester et de migrer immédiatement vos charges de travail vers la dernière version d'Amazon EMR. Pour plus d'informations, consultez le Blog sur la sécurité d'AWS.

Les versions Amazon EMR 2.x and 3.x, appelées versions AMI, sont disponibles pour les solutions pré-existantes qui en ont besoin pour des raisons de compatibilité. Il est déconseillé de créer de

nouveaux clusters ou de nouvelles solutions avec ces versions. Elles ne sont pas dotées des fonctions présentes dans les versions plus récentes et contiennent des packages d'application obsolètes.

Nous vous recommandons d'utiliser la version la plus récente d'Amazon EMR pour créer des solutions.

Il existe des différences significatives entre les versions 2.x et 3.x et les versions récentes d'Amazon EMR. Ces différences s'appliquent à différents domaines, comme la création et la configuration d'un cluster, ou encore les ports et la structure de répertoire des applications sur le cluster.

Cette section tente de répertorier les différences les plus importantes pour Amazon EMR, ainsi que les différences spécifiques portant sur la gestion et la configuration des applications. Elle n'est pas exhaustive. Si vous créez et utilisez des clusters dans les versions 2.x ou 3.x, il se peut que vous rencontriez des différences non répertoriées dans cette section.

Rubriques

- Création d'un cluster avec des versions AMI plus récentes d'Amazon EMR
- Installation d'applications avec des versions AMI antérieures d'Amazon EMR
- Personnalisation de la configuration des clusters et des applications avec des versions AMI antérieures d'Amazon EMR
- Caractéristiques de l'application Hive pour les versions AMI antérieures d'Amazon EMR
- HBase Caractéristiques de l'application pour les versions AMI antérieures d'Amazon EMR
- Caractéristiques de l'application Pig pour les versions AMI antérieures d'Amazon EMR
- Caractéristiques de l'application Spark pour les versions AMI antérieures d'Amazon EMR
- DistCp L'utilitaire S3 se différencie des versions AMI antérieures d'Amazon EMR

Création d'un cluster avec des versions AMI plus récentes d'Amazon EMR

Les versions Amazon EMR 2.x et 3.x sont référencées par la version AMI. À partir des versions 4.0.0 et ultérieures d'Amazon EMR, les publications sont référencées par version, à l'aide d'une étiquette de version telle que `emr-5.11.0`. Ce changement est particulièrement visible lorsque vous créez un cluster à l'aide du AWS CLI ou par programmation.

Lorsque vous utilisez le AWS CLI pour créer un cluster à l'aide d'une version de version AMI, utilisez l'`--ami-version`option, par exemple,`--ami-version 3.11.0`. De nombreuses options, fonctions et applications introduites dans les versions 4.0.0 et ultérieures d'Amazon EMR ne sont pas

disponibles lorsque vous spécifiez une option `--ami-version`. Pour plus d'informations, consultez create-cluster dans la Référence des commandes de la AWS CLI .

L'exemple de AWS CLI commande suivant lance un cluster à l'aide d'une version AMI.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.11.0 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,\
InstanceType=m3.xlarge InstanceGroupType=CORE,InstanceCount=2,\
InstanceType=m3.xlarge --bootstrap-actions Path=s3://elasticmapreduce/bootstrap-
actions/configure-hadoop,\
Name="Configuring infinite JVM reuse",Args=["-m","mapred.job.reuse.jvm.num.tasks=-1"]
```

Par programmation, toutes les versions d'Amazon EMR utilisent l'action `RunJobFlowRequest` de l'API EMR pour créer des clusters. Dans l'exemple suivant, le code Java crée un cluster à l'aide de la version AMI 3.11.0.

```
RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("AmiVersion Cluster")
   .withAmiVersion("3.11.0")
   .withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myKeyPair")
    .withInstanceCount(1)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m3.xlarge")
    .withSlaveInstanceType("m3.xlarge");
```

L'appel `RunJobFlowRequest` suivant utilise quant à lui une étiquette de version :

```
RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("ReleaseLabel Cluster")
   .withReleaseLabel("emr-7.9.0")
   .withInstances(new JobFlowInstancesConfig()
```

```
    .withEc2KeyName("myKeyPair")
    .withInstanceCount(1)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m3.xlarge")
    .withSlaveInstanceType("m3.xlarge");
```

Configuration de la taille du cluster

Lorsque votre cluster s'exécute, Hadoop détermine le nombre de tâches de mappeur et de réducteur nécessaires pour traiter les données. Les clusters les plus grands doivent avoir plus de tâches pour une meilleure utilisation des ressources et une plus courte durée de traitement. En règle générale, un cluster EMR conserve la même taille tout au long de sa vie ; vous définissez le nombre de tâches lorsque vous créez le cluster. Lorsque vous redimensionnez un cluster en cours d'exécution, vous pouvez modifier le traitement au cours de l'exécution du cluster. Par conséquent, au lieu d'utiliser un nombre fixe de tâches, vous pouvez modifier le nombre de tâches au cours de la vie du cluster. Il existe deux options de configuration qui vous aideront à définir le nombre idéal de tâches :

- `mapred.map.tasksperslot`
- `mapred.reduce.tasksperslot`

Vous pouvez définir ces deux options dans le fichier `mapred-conf.xml`. Lorsque vous soumettez un travail au cluster, le client du travail vérifie le nombre total actuel d'emplacements de mappage et de réduction disponibles sur l'ensemble du cluster. Le client du travail utilise alors les équations suivantes pour définir le nombre de tâches :

- `mapred.map.tasks = mapred.map.tasksperslot` * emplacements de mappage dans le cluster
- `mapred.reduce.tasks` = `mapred.reduce.tasksperslot` * emplacements de réduction dans le cluster

Le client du travail lit uniquement le paramètre `tasksperslot` si le nombre de tâches n'est pas configuré. Vous pouvez remplacer le nombre de tâches à tout moment, pour tous les clusters via une action d'amorçage ou individuellement, travail par travail, en ajoutant une étape pour modifier la configuration.

Amazon EMR supporte les défaillances des nœuds de tâches et continue l'exécution du cluster même si un nœud de tâches cesse d'être disponible. Amazon EMR alloue automatiquement des nœuds de tâches supplémentaires pour remplacer les nœuds défaillants.

Vous pouvez avoir un nombre différent de nœuds de tâches pour chaque étape du cluster. Vous pouvez également ajouter une étape à un cluster en cours d'exécution pour modifier le nombre de nœuds de tâches. Étant donné que toutes les étapes sont garanties de s'exécuter de manière séquentielle par défaut, vous pouvez spécifier le nombre de nœuds de tâches en cours d'exécution pour n'importe quelle étape.

Installation d'applications avec des versions AMI antérieures d'Amazon EMR

Lorsque vous utilisez une version AMI, les applications sont installées de différentes manières, notamment en utilisant le `NewSupportedProducts` paramètre de l'RunJobFlowaction, en utilisant les actions d'amorçage et en utilisant l'action Step.

Personnalisation de la configuration des clusters et des applications avec des versions AMI antérieures d'Amazon EMR

Amazon EMR version 4.0.0 a introduit une méthode simplifiée pour configurer des applications à l'aide de classifications de configuration. Pour de plus amples informations, veuillez consulter Configuration des applications. Lorsque vous utilisez une version AMI, vous configurez les applications à l'aide des actions d'amorçage et des arguments que vous transmettez. Par exemple, les actions d'amorçage `configure-hadoop` et `configure-daemons` définissent les propriétés des environnements Hadoop et YARN telles que `--namenode-heap-size`. Dans les versions plus récentes, ces propriétés sont configurées à l'aide des classifications de configuration `hadoop-env` et `yarn-env`. Pour les actions de bootstrap qui exécutent des configurations courantes, consultez le emr-bootstrap-actions référentiel sur Github.

Les tableaux suivants associent les actions d'amorçage aux classifications de configuration correspondantes dans les versions plus récentes d'Amazon EMR.

Hadoop

Nom du fichier d'application concerné	Action d'amorçage de la version d'AMI	Classification de configuration
`core-site.xml`	`configure-hadoop -c`	`core-site`
`log4j.properties`	`configure-hadoop -l`	`hadoop-log4j`
`hdfs-site.xml`	`configure-hadoop -s`	`hdfs-site`

Nom du fichier d'application concerné	Action d'amorçage de la version d'AMI	Classification de configuration
N/A	s/o	`hdfs-encryption-zones`
`mapred-site.xml`	`configure-hadoop -m`	`mapred-site`
`yarn-site.xml`	`configure-hadoop -y`	`yarn-site`
`httpfs-site.xml`	`configure-hadoop -t`	`httpfs-site`
`capacity-scheduler.xml`	`configure-hadoop -z`	`capacity-scheduler`
`yarn-env.sh`	`configure-daemons --resourcemanager-opts`	`yarn-env`

Hive

Nom du fichier d'application concerné	Action d'amorçage de la version d'AMI	Classification de configuration
`hive-env.sh`	N/A	`hive-env`
`hive-site.xml`	`hive-script --install -hive-site ${MY_HIVE_SITE_FILE}`	`hive-site`
`hive-exec-log4j.properties`	N/A	`hive-exec-log4j`
`hive-log4j.properties`	s/o	`hive-log4j`

EMRFS

Nom du fichier d'application concerné	Action d'amorçage de la version d'AMI	Classification de configuration
`emrfs-site.xml`	`configure-hadoop -e`	`emrfs-site`
N/A	`s3get -s s3://custom-provider.jar -d /usr/share/aws/emr/auxlib/`	`emrfs-site` (avec le nouveau paramètre `fs.s3.cse.encryptionMaterialsProvider.uri`)

Pour obtenir une liste de toutes les classifications, consultez Configuration des applications.

Variables d'environnement d'application

Lorsque vous utilisez une version AMI, un script `hadoop-user-env.sh` est utilisé avec l'action d'amorçage `configure-daemons` pour configurer l'environnement Hadoop. Les actions suivantes sont incluses dans le script :

```
#!/bin/bash
export HADOOP_USER_CLASSPATH_FIRST=true;
echo "HADOOP_CLASSPATH=/path/to/my.jar" >> /home/hadoop/conf/hadoop-user-env.sh
```

Dans Amazon EMR version 4.x, vous procédez de la même façon en utilisant la classification de configuration `hadoop-env`, comme illustrée dans l'exemple suivant :

```
[
      {
         "Classification":"hadoop-env",
         "Properties":{

         },
         "Configurations":[
            {
               "Classification":"export",
               "Properties":{
                  "HADOOP_USER_CLASSPATH_FIRST":"true",
                  "HADOOP_CLASSPATH":"/path/to/my.jar"
```

```
            }
         }
      ]
   }
]
```

Autre exemple : l'utilisation de `configure-daemons` et la transmission de `--namenode-heap-size=2048` et `--namenode-opts=-XX:GCTimeRatio=19` équivalent aux classifications de configuration suivantes.

```
[
      {
         "Classification":"hadoop-env",
         "Properties":{

         },
         "Configurations":[
            {
               "Classification":"export",
               "Properties":{
                  "HADOOP_DATANODE_HEAPSIZE":  "2048",
            "HADOOP_NAMENODE_OPTS":  "-XX:GCTimeRatio=19"
               }
            }
         ]
      }
   ]
```

D'autres variables de l'environnement d'application ne sont plus définies dans `/home/hadoop/.bashrc`. Au lieu de cela, elles sont principalement définies dans les fichiers `/etc/default` par composant ou par application, par exemple, `/etc/default/hadoop`. Les scripts wrapper `/usr/bin/` installés par l'application RPMs peuvent également définir des variables d'environnement supplémentaires avant d'impliquer le script bin lui-même.

Ports de service

Lorsque vous employez une version AMI, certains services utilisent des ports personnalisés.

Modifications des paramètres de port

Paramètre	Version AMI 3.x	Open source par défaut
fs.default.name	hdfs://emrDeterminedIP:9000	par défaut (hdfs : *emrDeterminedIP* //:8020)
dfs.datanode.address	0.0.0.0:9200	valeur par défaut (0.0.0.0:50010)
dfs.datanode.http.address	0.0.0.0:9102	valeur par défaut (0.0.0.0:50075)
dfs.datanode.https.address	0.0.0.0:9402	valeur par défaut (0.0.0.0:50475)
dfs.datanode.ipc.address	0.0.0.0:9201	valeur par défaut (0.0.0.0:50020)
dfs.http.address	0.0.0.0:9101	valeur par défaut (0.0.0.0:50070)
dfs.https.address	0.0.0.0:9202	valeur par défaut (0.0.0.0:50470)
dfs.secondary.http.address	0.0.0.0:9104	valeur par défaut (0.0.0.0:50090)
yarn.nodemanager.address	0.0.0.0:9103	valeur par défaut (${yarn.nodemanager.hostname}:0)
yarn.nodemanager.localizer.address	0.0.0.0:9033	valeur par défaut (${yarn.nodemanager.hostname}:8040)
yarn.nodemanager.webapp.address	0.0.0.0:9035	valeur par défaut (${yarn.nodemanager.hostname}:8042)
yarn.resourcemanager.address	*emrDeterminedIP* :9022	valeur par défaut (${yarn.resourcemanager.hostname}:8032)

Paramètre	Version AMI 3.x	Open source par défaut
yarn.resourcemanager.admin.address	*emrDeterminedIP* :9025	valeur par défaut (${yarn.resourcemanager.hostname}:8033)
yarn.resourcemanager.resource-tracker.address	*emrDeterminedIP* :9023	valeur par défaut (${yarn.resourcemanager.hostname}:8031)
yarn.resourcemanager.scheduler.address	*emrDeterminedIP* :9024	valeur par défaut (${yarn.resourcemanager.hostname}:8030)
yarn.resourcemanager.webapp.address	0.0.0.0:9026	valeur par défaut (${yarn.resourcemanager.hostname}:8088)
yarn.web-proxy.address	*emrDeterminedIP* :9046	valeur par défaut (no-value)
yarn.resourcemanager.hostname	0.0.0.0 (par défaut)	*emrDeterminedIP*

> **Note**
>
> *emrDeterminedIP*Il s'agit d'une adresse IP qui est générée par Amazon EMR.

Users

Lorsque vous utilisez une version AMI, l'utilisateur hadoop exécute tous les processus et possède tous les fichiers. Dans Amazon EMR 4.0.0 et versions ultérieures, les utilisateurs existent au niveau de l'application et des composants.

Séquence d'installation, artefacts installés et emplacement des fichiers journaux

Lorsque vous utilisez une version AMI, les artefacts de l'application et leurs répertoires de configuration sont installés dans le répertoire /home/hadoop/*application*. Par exemple, si

vous avez installé Hive, le répertoire est `/home/hadoop/hive`. A partir des versions 4.0.0 et ultérieures d'Amazon EMR, les artefacts de l'application sont installés dans le répertoire `/usr/lib/`*`application`*. Lorsque vous utilisez une version AMI, les fichiers journaux sont situés à différents emplacements. Le tableau ci-dessous répertorie ces emplacements.

Modification des emplacements des journaux sur Amazon S3

Démon ou application	Emplacement du répertoire
instance-state	noeud/ /instance-state/ *instance-id*
hadoop-hdfs-namenode	démons//.log *instance-id* hadoop-hadoop-namenode
hadoop-hdfs-datanode	démons//.log *instance-id* hadoop-hadoop-datanode
fil de hadoop () ResourceManager	démons//*instance-id* yarn-hadoop-resourcemanager
hadoop-yarn (serveur Proxy)	démons//*instance-id* yarn-hadoop-proxyserver
mapred-historyserver	démons//*instance-id*
httpfs	daemons/ /httpfs.log *instance-id*
hive-server	noeud/ /hive-server/hive-server.log *instance-id*
hive-metastore	noeud/ /apps/hive.log *instance-id*
Interface de ligne de commande Hive	noeud/ /apps/hive.log *instance-id*
Journaux utilisateur et journaux de conteneur des applications YARN	task-attempts/
Mahout	N/A
Pig	N/A

Démon ou application	Emplacement du répertoire
spark-historyserver	N/A
Fichiers d'historique des tâches mapreduce	jobs/

Exécuteur de commandes

Lorsque vous utilisez une version AMI, de nombreux scripts ou programmes comme `/home/hadoop/contrib/streaming/hadoop-streaming.jar` ne sont pas placés dans l'environnement du chemin d'accès de connexion au shell, ce qui signifie que vous devez spécifier le chemin d'accès complet lorsque vous utilisez un fichier jar (command-runner.jar ou script-runner.jar, par exemple) pour exécuter les scripts. Le fichier `command-runner.jar` se trouve sur l'image AMI, si bien qu'il n'est pas nécessaire de connaître un URI complet comme c'était le cas avec `script-runner.jar`.

Facteur de réplication

Le facteur de réplication vous permet de configurer à quel moment démarrer une machine virtuelle Java. Vous pouvez démarrer une nouvelle JVM Hadoop pour chaque tâche, ce qui permet une meilleure isolation des tâches, ou vous pouvez les partager JVMs entre les tâches, ce qui permet de réduire les frais de structure. Si vous traitez un grand nombre de petits fichiers, il est logique de réutiliser la JVM plusieurs fois pour amortir les coûts de démarrage. Toutefois, si chaque tâche prend du temps ou traite une grande quantité de données, vous pouvez choisir de ne pas réutiliser la JVM, afin de vous assurer que l'ensemble de la mémoire est libérée pour les tâches suivantes. Lorsque vous utilisez une version AMI, vous pouvez personnaliser le facteur de réplication à l'aide de l'action d'amorçage `configure-hadoop` pour définir la propriété `mapred.job.reuse.jvm.num.tasks`.

L'exemple suivant illustre la définition du facteur de réutilisation de la machine virtuelle Java pour une réutilisation infinie de cette dernière.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.11.0 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=m3.xlarge \
InstanceGroupType=CORE,InstanceCount=2,InstanceType=m3.xlarge \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hadoop,\
Name="Configuring infinite JVM reuse",Args=["-m","mapred.job.reuse.jvm.num.tasks=-1"]
```

Caractéristiques de l'application Hive pour les versions AMI antérieures d'Amazon EMR

Les fichiers journaux

Lorsque vous utilisez les versions AMI 2.x et 3.x d'Amazon EMR, les fichiers journaux Hive sont enregistrés dans `/mnt/var/log/apps/`. La version de Hive que vous exécutez détermine le nom du fichier journal, comme illustré dans le tableau suivant. Cela permet de prendre en charge les versions simultanées de Hive.

Version de Hive	Nom du fichier journal
0.13.1	`hive.log` **Note** A partir de cette version, Amazon EMR utilise un nom de fichier non versionné, `hive.log`. Les versions mineures partagent le même emplacement du journal que la version majeure.
0.11.0	`hive_0110.log` **Note** Les versions mineures de Hive 0.11.0, par exemple 0.11.0.1, partagent le même emplacement du fichier journal que Hive 0.11.0.

Version de Hive	Nom du fichier journal
0.8.1	`hive_081.log` **Note** Les versions mineures de Hive 0.8.1, par exemple Hive 0.8.1.1, partagent le même emplacement du fichier journal que Hive 0.8.1.
0.7.1	`hive_07_1.log` **Note** Les versions mineures de Hive 0.7.1, par exemple Hive 0.7.1.3 et Hive 0.7.1.4, partagent le même emplacement du fichier journal que Hive 0.7.1.
0.7	`hive_07.log`
0.5	`hive_05.log`
0.4	`hive.log`

Fonctionnalité de format d'entrée avec fractionnement

Pour implémenter la fonctionnalité de format d'entrée avec fractionnement à l'aide des versions de Hive antérieures à 0.13.1 (versions AMI d'Amazon EMR antérieures à 3.11.0), utilisez les commandes suivantes :

```
hive> set hive.input.format=org.apache.hadoop.hive.ql.io.HiveCombineSplitsInputFormat;
hive> set mapred.min.split.size=100000000;
```

Cette fonctionnalité est devenue obsolète avec Hive 0.13.1. Pour profiter de la même fonctionnalité de format d'entrée avec fractionnement dans la version 3.11.0 AMI Amazon EMR, procédez comme suit :

```
set hive.hadoop.supports.splittable.combineinputformat=true;
```

Ports de service Thrift

Thrift est une infrastructure RPC qui définit un format de sérialisation binaire compact utilisé pour conserver les structures de données afin de les analyser ultérieurement. Généralement, Hive configure le serveur pour qu'il fonctionne sur les ports suivants.

Version de Hive	Numéro de port
Hive 0.13.1	10 000
Hive 0.11.0	10004
Hive 0.8.1	10003
Hive 0.7.1	10002
Hive 0.7	10001
Hive 0.5	10 000

Pour plus d'informations sur les services thrift, consultez http://wiki.apache.org/thrift/.

Utilisation de Hive pour récupérer des partitions

Amazon EMR inclut une instruction du langage de requête Hive qui récupère les partitions d'une table depuis des données de table situées dans Amazon S3. L'exemple suivant en est une illustration.

```
CREATE EXTERNAL TABLE (json string) raw_impression
PARTITIONED BY (dt string)
LOCATION 's3://elastic-mapreduce/samples/hive-ads/tables/impressions';
ALTER TABLE logs RECOVER PARTITIONS;
```

Les données et les répertoires de la partition doit être à l'emplacement spécifié dans la définition de table et doivent être nommées selon la convention Hive : par exemple, `dt=2009-01-01`.

> **Note**
>
> Après Hive 0.13.1, cette fonctionnalité est prise en charge en mode natif à l'aide de `msck repair table` et par conséquent la prise en charge de `recover partitions` n'est plus assurée. Pour plus d'informations, consultez https://cwiki.apache. org/confluence/display/Hive/LanguageManual+DDL.

Transmission d'une variable Hive à un script

Pour transmettre une variable à une étape Hive à l' AWS CLI aide de la commande suivante, remplacez *myKey* par le nom de votre paire de EC2 clés et remplacez par le nom *amzn-s3-demo-bucket* de votre compartiment. Dans cet exemple, `SAMPLE` est une valeur de variable précédée par le commutateur `-d`. Cette variable est définie dans le script Hive comme suit : `${SAMPLE}`.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.9 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m3.xlarge --instance-count 3 \
--steps Type=Hive,Name="Hive Program",ActionOnFailure=CONTINUE,\
Args=[-f,s3://elasticmapreduce/samples/hive-ads/libs/response-time-stats.q,-d,\
INPUT=s3://elasticmapreduce/samples/hive-ads/tables,-d,OUTPUT=s3://amzn-s3-demo-bucket/
hive-ads/output/,\
-d,SAMPLE=s3://elasticmapreduce/samples/hive-ads/]
```

Spécification d'un emplacement de metastore externe

La procédure suivante vous montre comment remplacer les valeurs de configuration par défaut pour l'emplacement du metastore Hive et démarrer un cluster en utilisant l'emplacement du metastore reconfiguré.

Pour créer un metastore situé en dehors du cluster EMR

1. Créez une base de données MySQL ou Aurora à l'aide d'Amazon RDS.

   Pour plus d'informations sur la façon de créer une base de données Amazon RDS, consultez Mise en route sur Amazon RDS.

2. Modifiez vos groupes de sécurité pour autoriser les connexions JDBC entre votre base de données et le groupe de sécurité ElasticMapReduce-Master.

   Pour plus d'informations sur la façon de modifier vos groupes de sécurité pour l'accès, consultez Groupes de sécurité Amazon RDS dans le Guide de l'utilisateur Amazon RDS.

3. Définissez les valeurs de configuration JDBC dans `hive-site.xml` :

   a. Créez un fichier de configuration `hive-site.xml` contenant les éléments suivants :

```
<configuration>
  <property>
    <name>javax.jdo.option.ConnectionURL</name>
    <value>jdbc:mariadb://hostname:3306/hive?createDatabaseIfNotExist=true</
value>
    <description>JDBC connect string for a JDBC metastore</description>
  </property>
  <property>
    <name>javax.jdo.option.ConnectionUserName</name>
    <value>hive</value>
    <description>Username to use against metastore database</description>
  </property>
  <property>
    <name>javax.jdo.option.ConnectionPassword</name>
    <value>password</value>
    <description>Password to use against metastore database</description>
  </property>
</configuration>
```

*hostname*est l'adresse DNS de l'instance Amazon RDS exécutant la base de données. *username*et *password* sont les informations d'identification de votre base de données. Pour plus d'informations sur la connexion aux instances de bases de données MySQL et Aurora, consultez les sections Connexion à une instance de base de données exécutant le moteur de base de données MySQL et Connexion à un cluster de bases de données Aurora dans le Guide de l'utilisateur Amazon RDS.

Les pilotes JDBC sont installés par Amazon EMR.

> **Note**
>
> La propriété value ne doit pas contenir d'espaces ni de retours chariot. Elle doit figurer entièrement sur une seule ligne.

b. Enregistrez votre fichier `hive-site.xml` dans un emplacement sur Amazon S3, tel que `s3://amzn-s3-demo-bucket/hive-site.xml`.

4. Créez un cluster, en spécifiant l'emplacement Amazon S3 du fichier `hive-site.xml` personnalisé.

L'exemple de commande suivant illustre une AWS CLI commande qui effectue cette opération.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.10 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m3.xlarge --instance-count 3 \
--bootstrap-actions Name="Install Hive Site Configuration",\
Path="s3://region.elasticmapreduce/libs/hive/hive-script",\
Args=["--base-path","s3://elasticmapreduce/libs/hive","--install-hive-site",\
"--hive-site=s3://amzn-s3-demo-bucket/hive-site.xml","--hive-versions","latest"]
```

Connexion à Hive à l'aide de JDBC

Pour vous connecter à Hive via JDBC, vous devez télécharger le pilote JDBC et installer un client SQL. L'exemple suivant illustre l'utilisation de SQL Workbench/J pour vous connecter à Hive à l'aide de JDBC.

Pour télécharger les pilotes JDBC

1. Téléchargez et extrayez les pilotes correspondant aux versions de Hive auxquelles vous souhaitez accéder. La version de Hive varie en fonction de l'image AMI que vous choisissez lorsque vous créez un cluster Amazon EMR.

   - Pilotes JDBC Hive 0.13.1 : JDBC_1.0.4.1004.zip https://amazon-odbc-jdbc-drivers.s3.amazonaws.com/public/ AmazonHive
   - Pilotes JDBC Hive 0.11.0 : https://mvnrepository.com/artifact/org.apache.hive/hive-jdbc/0.11.0
   - Pilotes JDBC Hive 0.8.1 : https://mvnrepository.com/artifact/org.apache.hive/hive-jdbc/0.8.1

2. Installez SQL Workbench/J. Pour plus d'informations, consultez Installing and starting SQL Workbench/J dans le manuel de l'utilisateur SQL Workbench/J.
3. Créez un tunnel SSH vers le nœud maître du cluster. Le port de connexion diffère en fonction de la version de Hive. Les exemples fournis dans les tableaux ci-dessous concernent les utilisateurs de Linux (commandes `ssh`) et les commandes PuTTY pour les utilisateurs de Windows

   Commandes SSH Linux

Version de Hive	Command
0.13.1	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10000:localhost:10000 hadoop@` *`master-public-dns-name`*
0.11.0	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10004:localhost:10004 hadoop@` *`master-public-dns-name`*
0.8.1	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10003:localhost:10003 hadoop@` *`master-public-dns-name`*
0.7.1	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10002:localhost:10002 hadoop@` *`master-public-dns-name`*

Version de Hive	Command
0.7	`ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10001:localhost:10001 hadoop@ master-public-dns-name`
0.5	`ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10000:localhost:10000 hadoop@ master-public-dns-name`

Paramètres du tunnel PuTTY (Windows)

Version de Hive	Paramètres de tunnel
0.13.1	Port source : 10000 Destination : : *master-public-dns-name* 10000
0.11.0	Port source : 10004 Destination : 10004 *master-public-dns-name*
0.8.1	Port source : 10003 Destination : 10003 *master-public-dns-name*

4. Ajoutez le pilote JDBC à SQL Workbench.

   a. Dans la boîte de dialogue Select Connection Profile (Sélectionner le profil de connexion), sélectionnez Manage Drivers (Gérer les pilotes).

   b. Sélectionnez l'icône Create a new entry (Créer une nouvelle saisie) (page blanche).

   c. Dans le champ Name (Nom), saisissez **`Hive JDBC`**.

   d. Pour Library (Bibliothèque), cliquez sur l'icône Select the JAR file(s) (Sélectionner le(s) fichier(s) JAR).

   e. Sélectionnez les fichiers JAR comme indiqué dans le tableau suivant.

Version du pilote Hive	Fichiers JAR à ajouter

0.13.1	`hive_metastore.jar`
	`hive_service.jar`
	`HiveJDBC3.jar`

Version du pilote Hive	Fichiers JAR à ajouter
	libfb303-0.9.0.jar libthrift-0.9.0.jar log4j-1.2.14.jar ql.jar slf4j-api-1.5.8.jar slf4j-log4j12-1.5.8.jar TCLIServiceClient.jar
0.11.0	hadoop-core-1.0.3.jar hive-exec-0.11.0.jar hive-jdbc-0.11.0.jar hive-metastore-0.11.0.jar hive-service-0.11.0.jar libfb303-0.9.0.jar commons-logging-1.0.4.jar slf4j-api-1.6.1.jar
0.8.1	hadoop-core-0.20.205.jar hive-exec-0.8.1.jar hive-jdbc-0.8.1.jar hive-metastore-0.8.1.jar hive-service-0.8.1.jar libfb303-0.7.0.jar libthrift-0.7.0.jar log4j-1.2.15.jar slf4j-api-1.6.1.jar slf4j-log4j12-1.6.1.jar

Version du pilote Hive	Fichiers JAR à ajouter

0.7.1	`hadoop-0.20-core.jar`
	`hive-exec-0.7.1.jar`
	`hive-jdbc-0.7.1.jar`
	`hive-metastore-0.7.1.jar`
	`hive-service-0.7.1.jar`
	`libfb303.jar`
	`commons-logging-1.0.4.jar`
	`slf4j-api-1.6.1.jar`
	`slf4j-log4j12-1.6.1.jar`

0.7	`hadoop-0.20-core.jar`
	`hive-exec-0.7.0.jar`
	`hive-jdbc-0.7.0.jar`
	`hive-metastore-0.7.0.jar`
	`hive-service-0.7.0.jar`
	`libfb303.jar`
	`commons-logging-1.0.4.jar`
	`slf4j-api-1.5.6.jar`
	`slf4j-log4j12-1.5.6.jar`

0.5	`hadoop-0.20-core.jar`
	`hive-exec-0.5.0.jar`
	`hive-jdbc-0.5.0.jar`
	`hive-metastore-0.5.0.jar`
	`hive-service-0.5.0.jar`
	`libfb303.jar`
	`log4j-1.2.15.jar`
	`commons-logging-1.0.4.jar`

f. Dans la boîte de dialogue Please select one driver (Veuillez sélectionner un pilote), sélectionnez un pilote en vous conformant au tableau suivant, puis cliquez sur OK.

Version de Hive	Nom de classe du pilote
0.13.1	`com.amazon.hive.jdbc3.HS2Driver`
0.11.0	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.8.1	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.7.1	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.7	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.5	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`

5. Lorsque vous revenez à la boîte de dialogue Select Connection Profile (Sélectionner le profil de connexion), vérifiez que le champ Driver (Pilote) est défini sur Hive JDBC et fournissez la chaîne de connexion JDBC dans le champ URL, en vous conformant au tableau suivant.

Version de Hive	Chaîne de connexion JDBC
0.13.1	`jdbc:hive2://localhost:10000/default`
0.11.0	`jdbc:hive://localhost:10004/default`
0.8.1	`jdbc:hive://localhost:10003/default`

Si votre cluster utilise la version 3.3.1 d'AMI ou une version ultérieure, dans la boîte de dialogue Select Connection Profile (Sélectionner le profil de connexion), tapez **hadoop** dans le champ Username (Nom d'utilisateur).

HBase Caractéristiques de l'application pour les versions AMI antérieures d'Amazon EMR

HBase Versions prises en charge

HBase version	Version d'AMI	AWS CLI paramètres de configuration	HBase détails de la version

0.94.18	3.1.0 et ultérieures	`--ami-version 3.1`
		`--ami-version 3.2`
		`--ami-version 3.3`
		`--applications Name=HBase`
			• Correctifs de bogues et améliorations

0.94.7	3.0-3.0.4	`--ami-version 3.0` `--applications Name=HBase`	
0.92	2.2 et ultérieures	`--ami-version 2.2 or later` `--applications Name=HBase`	

HBase Prérequis des clusters

Un cluster créé à l'aide des versions AMI 2.x et 3.x d'Amazon EMR doit respecter les conditions requises suivantes pour. HBase

- L' AWS CLI (facultatif) — Pour interagir à HBase l'aide de la ligne de commande, téléchargez et installez la dernière version de l' AWS CLI. Pour plus d'informations, consultez Installation d' AWS Command Line Interface dans le Guide de l'utilisateur AWS Command Line Interface .
- Au moins deux instances (facultatif) : le nœud principal du cluster exécute le serveur HBase maître et Zookeeper, et les nœuds de tâche exécutent les serveurs de HBase région. Pour de meilleures performances, les HBase clusters doivent s'exécuter sur au moins deux EC2 instances, mais vous pouvez les exécuter HBase sur un seul nœud à des fins d'évaluation.
- Cluster de longue durée, s'exécute HBase uniquement sur des clusters de longue durée. Par défaut, l'interface de ligne de commande et la console Amazon EMR créent des clusters de longue durée.
- Un jeu de paires de EC2 clés Amazon (recommandée) : pour utiliser le protocole Secure Shell (SSH) pour vous connecter avec le nœud principal et exécuter les commandes de HBase shell, vous devez utiliser une paire de EC2 clés Amazon lorsque vous créez le cluster.
- Versions correctes d'AMI et de Hadoop — Les HBase clusters sont actuellement pris en charge uniquement sur Hadoop 20.205 ou version ultérieure.
- Ganglia (facultatif) — Pour surveiller les métriques de HBase performance, installez Ganglia lorsque vous créez le cluster.
- Un compartiment Amazon S3 pour les journaux (facultatif) — Les journaux de HBase sont disponibles sur le nœud principal. Si vous voulez que ces journaux soient copiés dans Amazon S3, spécifiez un compartiment S3 qui contiendra les fichiers journaux lorsque vous créez le cluster.

Création d'un cluster avec HBase

Le tableau suivant répertorie les options qui sont disponibles lorsque vous utilisez la console pour créer un cluster à HBase l'aide d'une version AMI Amazon EMR.

Champ	Action
Restore from backup (Rétablir à partir d'une sauvegarde)	Spécifiez s'il convient de précharger le HBase cluster avec les données stockées dans Amazon S3.

Champ	Action
Backup location (Emplacement de sauvegarde)	Spécifiez l'URI où la sauvegarde à restaurer réside dans Amazon S3.
Backup version (Version de sauvegarde)	Éventuellement, spécifiez le nom de version de la sauvegarde dans Backup Location (Emplacement de sauvegarde) à utiliser. Si vous laissez ce champ vide, Amazon EMR utilise la dernière sauvegarde dans Emplacement de sauvegarde pour renseigner le nouveau cluster. HBase
Schedule Regular Backups (Planifier des sauvegardes régulières)	Spécifiez s'il convient de planifier des sauvegardes incrémentielles automatiq ues. La première sauvegarde est une sauvegarde complète pour créer une référence pour les futures sauvegardes incrémentielles.
Consisten t backup (Sauvegarde uniforme)	Spécifiez si les sauvegardes doivent être cohérentes. Une sauvegarde cohérente est une sauvegarde qui interrompt les opérations d'écriture au cours de la phase initiale de sauvegarde, pour la synchronisation entre les nœuds. Les éventuelles opérations d'écriture ainsi interrompues sont placées dans une file d'attente et reprises lorsque la synchronisation est terminée.
Backup frequency (Fréquence de sauvegarde)	Nombre de sauvegardes planifiées days/hours/minutes entre deux.
Backup location (Emplacement de sauvegarde)	L'URI Amazon S3 où les sauvegardes sont stockées. L'emplacement de sauvegarde de chaque HBase cluster doit être différent pour garantir que les sauvegardes différentielles restent correctes.
Backup start time (Heure de début de sauvegarde)	Spécifiez quand la première sauvegarde doit avoir lieu. Vous pouvez spécifier now, ce qui entraîne le commencement de la première sauvegarde lorsque le cluster est en cours d'exécution, ou saisir une date et une heure dans ISO format. Par exemple, 2012-06-15T20:00Z définit la date et l'heure de début sur le 15 juin 2012 à 20 h 00 UTC.

L'exemple de AWS CLI commande suivant lance un cluster avec HBase d'autres applications :

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
               --applications Name=Hue Name=Hive Name=Pig Name=HBase \
               --use-default-roles --ec2-attributes KeyName=myKey \
               --instance-type c1.xlarge --instance-count 3 --termination-protected
```

Une fois la connexion établie entre le Hive et les HBase clusters (comme indiqué dans la procédure précédente), vous pouvez accéder aux données stockées sur le HBase cluster en créant une table externe dans Hive.

L'exemple suivant, lorsqu'il est exécuté à partir de l'invite Hive, crée une table externe qui fait référence aux données stockées dans une HBase table appeléeinputTable. Vous pouvez ensuite faire référence `inputTable` dans les instructions Hive pour interroger et modifier les données stockées dans le HBase cluster.

> **Note**
>
> L'exemple suivant utilise protobuf-java-2.4.0a.jar dans AMI 2.3.3, mais vous devez modifier l'exemple pour qu'il corresponde à votre version. Pour vérifier quelle version du fichier JAR de tampons de protocole vous avez, exécutez la commande à l'invite de commande Hive : `! ls /home/hadoop/lib;`.

```
add jar lib/emr-metrics-1.0.jar ;
               add jar lib/protobuf-java-2.4.0a.jar ;

               set hbase.zookeeper.quorum=ec2-107-21-163-157.compute-1.amazonaws.com ;

               create external table inputTable (key string, value string)
                    stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
                     with serdeproperties ("hbase.columns.mapping" = ":key,f1:col1")
```

```
                    tblproperties ("hbase.table.name" = "t1");

                select count(*) from inputTable ;
```

Personnalisation de la configuration HBase

Bien que les paramètres par défaut soient adaptés à la plupart des applications, vous pouvez modifier vos paramètres HBase de configuration. Pour ce faire, exécutez l'un des deux scripts d'action d'amorçage :

- configure-hbase-daemons— Configure les propriétés des démons principal, regionserver et zookeeper. Ces propriétés incluent la taille du segment de mémoire et les options à transmettre à la machine virtuelle Java (JVM) au démarrage du HBase démon. Vous définissez ces propriétés en tant qu'arguments dans l'action d'amorçage. Cette action d'amorçage modifie le fichier de configuration/home/hadoop/conf/hbase-user-env.sh sur le HBase cluster.
- configure-hbase — Configure les paramètres HBase spécifiques au site, tels que le port auquel le HBase maître doit se lier et le nombre maximal de fois où le client d'interface de ligne de commande client doit recommencer une action. Vous pouvez définir ces paramètres one-by-one, en tant qu'arguments dans l'action d'amorçage, ou vous pouvez spécifier l'emplacement d'un fichier de configuration XML dans Amazon S3. Cette action d'amorçage modifie le fichier de configuration the /home/hadoop/conf/hbase -site.xml sur le HBase cluster.

> **Note**
>
> Ces scripts, comme les autres actions d'amorçage, ne peuvent être exécutés que lorsque le cluster est créé ; vous ne pouvez pas les utiliser pour modifier la configuration d'un HBase cluster en cours d'exécution.

Lorsque vous exécutez les actions configure-hbase ou configure-hbase-daemonsbootstrap, les valeurs que vous spécifiez remplacent les valeurs par défaut. Toutes les valeurs que vous ne définissez pas explicitement reçoivent les valeurs par défaut.

La configuration HBase avec ces actions d'amorçage est analogue à l'utilisation d'actions d'amorçage dans Amazon EMR pour configurer les paramètres Hadoop et les propriétés de démon Hadoop. La différence est qu'il HBase n'y a pas d'options de mémoire par processus. Au lieu de cela, les options de mémoire sont définies à l'aide de l'`--`*`daemon`*`-opts`argument, qui *`daemon`* est remplacé par le nom du démon à configurer.

Configuration des HBase démons

Amazon EMR fournit une action d'amorçage que vous pouvez utiliser pour modifier la configuration des HBase démons, où se *region* trouve la région où vous lancez votre cluster. `s3://region.elasticmapreduce/bootstrap-actions/configure-hbase-daemons` HBase

Pour configurer HBase des démons à l'aide de AWS CLI, ajoutez l'action `configure-hbase-daemons` bootstrap lorsque vous lancez le cluster afin de configurer un ou plusieurs démons. HBase Vous pouvez définir les propriétés suivantes :

Propriété	Description
`hbase-master-opts`	Options qui contrôlent la façon dont la machine virtuelle Java exécute le démon maître. Si elles sont définies, elles remplacent les variables HBASE_MASTER_OPTS par défaut.
`regionserver-opts`	Options qui contrôlent la façon dont la machine virtuelle Java exécute le démon de serveur de région. Si elles sont définies, elles remplacent les variables HBASE_REGIONSERVER_OPTS par défaut.
`zookeeper-opts`	Options qui contrôlent la façon dont la machine virtuelle Java exécute le démon zookeeper. Si elles sont définies, elles remplacent les variables HBASE_ZOOKEEPER_OPTS par défaut.

Pour plus d'informations sur ces options, consultez le fichier hbase-env.sh dans la HBase documentation.

Une action d'amorçage permettant de configurer les valeurs de `zookeeper-opts` et `hbase-master-opts` est illustrée dans l'exemple suivant.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
--applications Name=Hue Name=Hive Name=Pig Name=HBase \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type c1.xlarge --instance-count 3 --termination-protected \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hbase-daemons,\
Args=["--hbase-zookeeper-opts=-Xmx1024m -XX:GCTimeRatio=19","--hbase-master-opts=-Xmx2048m","--hbase-regionserver-opts=-Xmx4096m"]
```

Configuration des paramètres de HBase site

Amazon EMR fournit une action d'amorçage que vous pouvez utiliser pour modifier la configuration de. `s3://elasticmapreduce/bootstrap-actions/configure-hbase` HBase Vous pouvez définir des valeurs de configuration one-by-one, en tant qu'arguments dans l'action d'amorçage, ou vous pouvez spécifier l'emplacement d'un fichier de configuration XML dans Amazon S3. La définition de valeurs de configuration one-by-one est utile si vous n'avez besoin de définir que quelques paramètres de configuration. Leur définition à l'aide d'un fichier XML est utile si vous avez beaucoup de modifications à apporter, ou si vous souhaitez enregistrer vos paramètres de configuration en vue de les réutiliser.

> **Note**
>
> Vous pouvez préfixer le nom de compartiment Amazon S3 avec un préfixe de régions`s3://region.elasticmapreduce/bootstrap-actions/configure-hbase`, tel que où se *`region`* trouve la région où vous lancez votre HBase cluster.

Cette action d'amorçage modifie le fichier `/home/hadoop/conf/hbase-site.xml` de configuration du HBase cluster. L'action bootstrap ne peut être exécutée que lorsque le HBase cluster est lancé.

Pour plus d'informations sur les paramètres de HBase site que vous pouvez configurer, veuillez consulter Configuration par défaut dans la HBase documentation concernant.

Définissez l'action `configure-hbase` bootstrap lorsque vous lancez le HBase cluster et spécifiez les valeurs `hbase-site.xml` à modifier.

Pour définir les paramètres individuels HBase du site à l'aide du AWS CLI

- Pour modifier le `hbase.hregion.max.filesize` paramètre, tapez la commande suivante et remplacez-la *myKey* par le nom de votre paire de EC2 clés Amazon.

  > **Note**
  >
  > Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
--applications Name=Hue Name=Hive Name=Pig Name=HBase \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type c1.xlarge --instance-count 3 --termination-protected \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-
hbase,Args=["-s","hbase.hregion.max.filesize=52428800"]
```

Pour définir les paramètres HBase du site à l'aide d'un fichier XML à l'aide du AWS CLI

1. Créez une version personnalisée de `hbase-site.xml`. Votre fichier personnalisé doit être un fichier XML valide. Pour réduire le risque d'introduire des erreurs, commencez par la copie par défaut de`hbase-site.xml`, située sur le nœud HBase principal Amazon EMR à l'adresse/ `home/hadoop/conf/hbase-site.xml`, puis modifiez une copie de ce fichier au lieu de créer un fichier complètement neuf. Vous pouvez donner un nouveau nom à votre nouveau fichier ou conserver le nom `hbase-site.xml`.
2. Chargez votre fichier `hbase-site.xml` personnalisé sur un compartiment Amazon S3. Il doit disposer d'autorisations pour que le AWS compte qui lance le cluster puisse accéder à ce fichier. Si le AWS compte qui lance le cluster est également propriétaire du compartiment Amazon S3, il bénéficie d'un accès.

3. Définissez l'action de démarrage configure-hbase lorsque vous lancez le HBase cluster et incluez l'emplacement de votre fichier personnalisé. `hbase-site.xml` L'exemple suivant définit les valeurs de configuration du HBase site selon celles spécifiées dans le fichiers`3://amzn-s3-demo-bucket/my-hbase-site.xml`. Tapez la commande suivante, *myKey* remplacez-la par le nom de votre paire de EC2 clés et remplacez *amzn-s3-demo-bucket* par le nom de votre compartiment Amazon S3.

   > **Note**
   >
   > Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

   ```
   aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
           --applications Name=Hue Name=Hive Name=Pig Name=HBase \
           --use-default-roles --ec2-attributes KeyName=myKey \
           --instance-type c1.xlarge --instance-count 3 --termination-protected \
           --bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-
   hbase,Args=["--site-config-file","s3://amzn-s3-demo-bucket/config.xml"]
   ```

   Si vous spécifiez plusieurs options pour personnaliser le HBase fonctionnement, vous devez ajouter un commutateur d'`-s`options à chaque paire clé-valeur, comme indiqué dans l'exemple suivant :

   ```
         --bootstrap-actions s3://elasticmapreduce/bootstrap-actions/configure-
   hbase,Args=["-s","zookeeper.session.timeout=60000"]
   ```

Une fois le proxy défini et la connexion SSH ouverte, vous pouvez afficher l' HBase interface utilisateur en ouvrant une fenêtre de navigateur avec http : //:60010/master-status*master-public-dns-name*, où se trouve *master-public-dns-name* l'adresse DNS publique du nœud maître du cluster. HBase

Vous pouvez consulter les HBase journaux actuels en utilisant SSH pour vous connecter au nœud principal et en accédant au `mnt/var/log/hbase` répertoire. Ces journaux ne sont pas disponibles

une fois que le cluster a été arrêté sauf si vous activez la journalisation dans Amazon S3 lorsque le cluster est lancé.

Sauvegarde et restauration HBase

Amazon EMR offre la possibilité de sauvegarder vos HBase données sur Amazon S3, manuellement ou selon une planification automatique. Vous pouvez effectuer des sauvegardes complètes et incrémentielles. Une fois que vous disposez d'une version sauvegardée des HBase données, vous pouvez restaurer cette version dans un HBase cluster. Vous pouvez effectuer une restauration sur un HBase cluster en cours d'exécution ou lancer un nouveau cluster prérempli de données sauvegardées.

Pendant le processus de sauvegarde, HBase continue d'exécuter les commandes d'écriture. Cela garantit la disponibilité du cluster tout au long de la sauvegarde, mais cela entraîne un risque d'incohérence entre les données en cours de sauvegarde et les éventuelles opérations d'écriture en cours d'exécution en parallèle. Pour comprendre les incohérences qui peuvent survenir, vous devez tenir compte du fait que les opérations d' HBase écriture sont réparties entre les nœuds de son cluster. Si une opération d'écriture se produit après qu'un nœud particulier a été interrogé, ces données ne sont pas incluses dans l'archive de sauvegarde. Vous pouvez même constater que les écritures antérieures sur le HBase cluster (envoyées à un nœud qui a déjà été interrogé) ne figurent peut-être pas dans l'archive de sauvegarde, alors que les écritures ultérieures (envoyées à un nœud avant qu'il ne soit interrogé) sont incluses.

Si une sauvegarde cohérente est requise, vous devez suspendre les écritures HBase pendant la partie initiale du processus de sauvegarde, c'est-à-dire la synchronisation entre les nœuds. Pour cela, vous pouvez spécifier le paramètre `--consistent` lorsque vous demandez une sauvegarde. Avec ce paramètre, les écritures intervenant au cours de cette période sont mises en file d'attente et exécutées dès que la synchronisation se termine. Vous pouvez également planifier des sauvegardes récurrentes, ce qui résout les incohérences au fil du temps, car les données qui manquent dans un cycle de sauvegarde sont sauvegardées lors du cycle suivant.

Lorsque vous sauvegardez HBase des données, vous devez spécifier un répertoire de sauvegarde différent pour chaque cluster. Un moyen simple pour cela consiste à utiliser l'identifiant du cluster dans le chemin d'accès spécifié pour le répertoire de sauvegarde. Par exemple, `s3://amzn-s3-demo-bucket/backups/j-3AEXXXXXX16F2`. Cela garantit que toutes les futures sauvegardes incrémentielles font référence au HBase cluster approprié.

Lorsque vous êtes prêt à supprimer les anciens fichiers de sauvegarde devenus inutiles, nous vous recommandons d'effectuer d'abord une sauvegarde complète de vos HBase données. Ceci garantit

la conservation de toutes les données et fournit un point de départ pour les futures sauvegardes incrémentielles. Une fois la sauvegarde complète terminée, vous pouvez accédez à l'emplacement de sauvegarde et supprimer manuellement les anciens fichiers de sauvegarde.

Le processus HBase de sauvegarde utilise S3 DistCp pour l'opération de copie, qui comporte certaines limites concernant l'espace de stockage temporaire des fichiers.

Sauvegarde et restauration à HBase l'aide de la console

La console permet de lancer un nouveau cluster et de le remplir avec les données d'une HBase sauvegarde précédente. Il vous permet également de planifier des sauvegardes incrémentielles périodiques des HBase données. Des fonctionnalités de sauvegarde et de restauration supplémentaires sont disponibles via l'interface de ligne de commande, telles que la possibilité de restaurer les données dans un cluster déjà en cours d'exécution, d'effectuer des sauvegardes manuelles et de planifier des sauvegardes complètes automatisées.

Pour remplir un nouveau cluster avec des HBase données archivées à l'aide de la console

1. Accédez à la nouvelle console Amazon EMR et sélectionnez Changer pour l'ancienne console depuis le menu latéral. Pour plus d'informations sur ce qu'implique le passage à l'ancienne console, consultez la rubrique Utilisation de l'ancienne console.
2. Choisissez Créer un cluster.
3. Dans la section Configuration logicielle, pour Applications supplémentaires, sélectionnez Configurer HBaseet ajouter.
4. Dans la boîte de dialogue Add Application (Ajouter une application), cochez Restore From Backup (Restituer à partir d'une sauvegarde).
5. Pour Backup Location, spécifiez l'emplacement de la sauvegarde à charger dans le nouveau HBase cluster. Il doit s'agir d'une URL Amazon S3 de la forme `s3://amzn-s3-demo-bucket/backups/`.
6. Pour Backup Version (Version de sauvegarde), vous avez la possibilité de spécifier le nom d'une version de sauvegarde à charger en définissant une valeur. Si vous ne définissez pas de valeur pour Version de sauvegarde, Amazon EMR charge la dernière sauvegarde à l'emplacement spécifié.
7. Choisissez Ajouter, puis créez le cluster avec d'autres options, si vous le souhaitez.

Pour planifier des sauvegardes automatisées des HBase données à l'aide de la console

1. Dans la section Configuration logicielle, pour Applications supplémentaires, sélectionnez Configurer HBaseet ajouter.
2. Choisissez Schedule Regular Backups (Planifier des sauvegardes régulières).
3. Spécifiez si les sauvegardes doivent être cohérentes. Une sauvegarde cohérente est une sauvegarde qui interrompt les opérations d'écriture au cours de la phase initiale de sauvegarde, pour la synchronisation entre les nœuds. Les éventuelles opérations d'écriture ainsi interrompues sont placées dans une file d'attente et reprises lorsque la synchronisation est terminée.
4. Définissez la fréquence à laquelle les sauvegardes doivent intervenir en entrant un nombre pour Backup Frequency (Fréquence de sauvegarde) et en choisissant Jours, Heures ou Minutes. La première sauvegarde automatisée qui s'exécute est une sauvegarde complète. Après cela, Amazon EMR enregistre des sauvegardes incrémentielles basées sur la planification que vous spécifiez.
5. Spécifiez l'emplacement dans Amazon S3 où les sauvegardes doivent être stockées. Chaque HBase cluster doit être sauvegardé dans un emplacement distinct dans Amazon S3, afin de garantir que les sauvegardes incrémentielles sont calculées correctement.
6. Spécifiez quand la première sauvegarde doit intervenir en définissant une valeur pour Backup Start Time (Heure de début de sauvegarde). Vous pouvez spécifier `now`, ce qui entraîne le commencement de la première sauvegarde lorsque le cluster est en cours d'exécution, ou saisir une date et une heure dans ISO format. Par exemple, 2013-09-26T20:00Z, définit la date et l'heure de début sur 26 septembre 2013 à 20 h 00 UTC.
7. Choisissez Ajouter.
8. Procédez à la création du cluster avec d'autres options, si vous le souhaitez.

Moniteur HBase avec CloudWatch

Amazon EMR signale trois métriques CloudWatch que vous pouvez utiliser pour surveiller vos HBase sauvegardes. Ces métriques sont CloudWatch poussées toutes les cinq minutes et sont fournies gratuitement.

Métrique	Description
`HBaseBackupFailed`	Si la dernière sauvegarde a échoué. La valeur est définie sur 0 par défaut et mise à jour sur 1 en cas d'échec de la

Métrique	Description
	tentative de sauvegarde précédente. Cette métrique n'est signalée que pour les HBase clusters. Cas d'utilisation : surveiller les HBase sauvegardes Unités : nombre
`HBaseMostRecentBackupDuration`	Délai nécessaire à l'exécution de la précédente sauvegarde. Cette métrique est définie même si la dernière sauvegarde a réussi ou a échoué. Lorsque la sauvegarde est en cours, cette métrique retourne le nombre de minutes qui se sont écoulées depuis le démarrage de la sauvegarde. Cette métrique n'est signalée que pour les HBase clusters. Cas d'utilisation : surveiller les HBase sauvegardes Unités : minutes
`HBaseTimeSinceLastSuccessfulBackup`	Nombre de minutes écoulées après le démarrage de la dernière HBase sauvegarde réussie sur votre cluster. Cette métrique n'est signalée que pour les HBase clusters. Cas d'utilisation : surveiller les HBase sauvegardes Unités : minutes

Configurer Ganglia pour HBase

Vous configurez Ganglia pour HBase utiliser l'action configure-hbase-for-gangliabootstrap. Cette action bootstrap est configurée HBase pour publier des métriques sur Ganglia.

Vous devez configurer HBase et Ganglia lorsque vous lancez le cluster ; les rapports Ganglia ne peuvent pas être ajoutés à un cluster en cours d'exécution.

Ganglia stocke également les fichiers journaux sur le serveur à l'adresse `/mnt/var/log/ganglia/rrds`. Si vous avez configuré votre cluster pour conserver les fichiers journaux dans un compartiment Amazon S3, les fichiers journaux Ganglia y sont conservés également.

Pour lancer un cluster avec Ganglia for HBase, utilisez l'action configure-hbase-for-gangliabootstrap comme indiqué dans l'exemple suivant.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
--applications Name=Hue Name=Hive Name=Pig Name=HBase Name=Ganglia \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type c1.xlarge --instance-count 3 --termination-protected \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hbase-for-ganglia
```

Une fois que le cluster a été lancé et Ganglia configuré, vous pouvez accéder aux graphiques et rapports Ganglia à l'aide de l'interface graphique en cours d'exécution sur le nœud maître.

Caractéristiques de l'application Pig pour les versions AMI antérieures d'Amazon EMR

Versions de Pig prises en charge

La version de Pig que vous pouvez ajouter à votre cluster dépend de la version d'AMI Amazon EMR et de la version de Hadoop que vous utilisez. Le tableau ci-dessous montre les versions d'AMI et les versions de Hadoop compatibles avec les différentes versions de Pig. Nous recommandons d'utiliser la dernière version de Pig disponible pour tirer parti des améliorations de performances et des nouvelles fonctionnalités.

Lorsque vous utilisez l'API pour installer Pig, la version par défaut est utilisée, sauf si vous spécifiez `--pig-versions` comme argument l'étape qui charge Pig sur le cluster lors de l'appel à RunJobFlow.

Version de Pig	Version d'AMI	Paramètres de configuration	Détails de la version de Pig

0.12.0
 	
Notes de mise à jour
 	
Documentation
	3.1.0 et ultérieures	`--ami-version 3.1`
		
		`--ami-version 3.2`
		
		`--ami-version 3.3`
			Ajoute la prise en charge des éléments suivants :
			
			• Streaming UDFs sans implément ations de JVM
			• Opérateurs ASSERT et IN
			• Expression CASE
			• AvroStorage en tant que fonction intégrée à Pig.
			• ParquetLoader et ParquetStorer en tant que fonctions intégrées
			• BigInteger et BigDecimal types

0.11.1.1
 	
Notes de mise à jour
 	
Documentation
	2.2 et ultérieures	`--pig-versions 0.11.1.1`
		
		`--ami-version 2.2`
			Améliore les performances de la commande LOAD, PigStorage si les données d'entrée se trouvent dans Amazon S3.

0.11.1
 	
Notes de mise à jour
 	
Documentation
	2.2 et ultérieures	`--pig-versions 0.11.1`
		
		`--ami-version 2.2`
			Ajoute la prise en charge de JDK 7, Hadoop 2, des fonctions Groovy définies par l'utilisa

Version de Pig	Version d'AMI	Paramètres de configuration	Détails de la version de Pig
			teur, de l' SchemaTuple optimisation, de nouveaux opérateurs, etc. Pour plus d'informations, consultez Journal de modifications Pig 0.11.1.
0.9.2.2 Notes de mise à jour Documentation	2.2 et ultérieures	`--pig-versions 0.9.2.2` `--ami-version 2.2`	Ajoute une prise en charge pour Hadoop 1.0.3.
0.9.2.1 Notes de mise à jour Documentation	2.2 et ultérieures	`--pig-versions 0.9.2.1` `--ami-version 2.2`	Ajoute la prise en charge de MapR.
0.9.2 Notes de mise à jour Documentation	2.2 et ultérieures	`--pig-versions 0.9.2` `--ami-version 2.2`	Inclut plusieurs améliorations de performances et correctifs. Pour obtenir des informations détaillées sur les modifications apportées dans la version Pig 0.9.2, consultez Journal de modifications Pig 0.9.2.

Version de Pig	Version d'AMI	Paramètres de configuration	Détails de la version de Pig

0.9.1
 	
Notes de mise à jour
 	
Documentation
	2.0	`--pig-versions 0.9.1`
		
		`--ami-version 2.0`	

0.6
 	
Notes de mise à jour
	1.0	`--pig-versions 0.6`
		
		`--ami-version 1.0`	

0.3
 	
Notes de mise à jour
	1.0	`--pig-versions 0.3`
		
		`--ami-version 1.0`	

Détails de la version de Pig

Amazon EMR prend en charge certaines versions de Pig auxquelles des correctifs Amazon EMR supplémentaires sont appliqués. Vous pouvez configurer la version de Pig à exécuter sur les clusters Amazon EMR. Pour plus d'informations sur cette étape, consultez Apache Pig. Les sections suivantes décrivent les différentes versions de Pig et les correctifs appliqués aux versions chargées sur Amazon EMR.

Correctifs Pig

Cette section décrit les correctifs personnalisés appliqués aux versions de Pig disponibles avec Amazon EMR.

Correctifs Pig 0.11.1.1

La version Amazon EMR de Pig 0.11.1.1 est une version de maintenance qui améliore les performances de la commande LOAD, PigStorage si les données d'entrée se trouvent dans Amazon S3.

Correctifs Pig 0.11.1

La version Amazon EMR de Pig 0.11.1 contient toutes les mises à jour fournies par l'Apache Software Foundation et les correctifs Amazon EMR cumulés depuis la version Pig 0.9.2.2. Il n'y a cependant pas de nouveaux correctifs spécifiques à Amazon EMR dans Pig 0.11.1.

Correctifs Pig 0.9.2

Apache Pig 0.9.2 est une version de maintenance de Pig. L'équipe Amazon EMR a appliqué les correctifs suivants à la version Amazon EMR de Pig 0.9.2.

Correctif	Description
PIG-1429	Ajout du type de données Boolean à Pig en tant que type de données de première classe. Pour plus d'informations, consultez https://issues.apache. org/jira/browse/PIG-1429. Statut : validé Correction dans la version Apache Pig : 0.10
PIG-1824	Prise en charge des modules d'importation dans Jython UDF. Pour plus d'informations, consultez https://issues.apache. org/jira/browse/PIG-1824. Statut : validé Correction dans la version Apache Pig : 0.10
PIG-2010	JARs Bundle enregistré dans le cache distribué. Pour plus d'informations, consultez https://issues.apache. org/jira/browse/PIG-2010. Statut : validé Correction dans la version Apache Pig : 0.11
PIG-2456	Ajout d'un fichier ~/.pigbootup dans lequel l'utilisateur peut spécifier des déclarations Pig par défaut. Pour plus d'informations, consultez https://issues.apache. org/jira/browse/PIG-2456.

Correctif	Description
	Statut : validé Correction dans la version Apache Pig : 0.11
PIG-2623	Support d'utilisation des chemins vers Amazon S3 pour l'enregis trement UDFs. Pour plus d'informations, consultez https://i ssues.apache. org/jira/browse/PIG-2623. Statut : validé Correction dans les versions Apache Pig : 0.10, 0.11

Correctifs Pig 0.9.1

L'équipe Amazon EMR a appliqué les correctifs suivants à la version Amazon EMR de Pig 0.9.1.

Correctif	Description
Prise en charge des fichiers JAR et des scripts Pig dans les systèmes de fichiers distribués	Ajout de la prise en charge de l'exécution de scripts et de l'enregistrement de fichiers JAR stockés dans HDFS, Amazon S3 ou d'autres systèmes de fichiers distribués. Pour plus d'informations, consultez https://issues.apache. org/jira/browse/ PIG-1505. Statut : validé Correction dans la version Apache Pig : 0.8.0
Prise en charge de plusieurs systèmes de fichiers dans Pig	Ajout de la prise en charge des scripts Pig pour lire les données d'un système de fichiers et les écrire dans un autre. Pour plus d'informations, consultez https://issues.apache. org/jira/browse/ PIG-1564. Statut : non validé Correction dans la version Apache Pig : non applicable

Correctif	Description
Ajouter la date, l'heure et la chaîne de caractères de Piggybank UDFs	Ajoutez la date, l'heure et la chaîne UDFs pour prendre en charge les scripts Pig personnalisés. Pour plus d'informations, consultez https://issues.apache. org/jira/browse/PIG-1565. Statut : non validé Correction dans la version Apache Pig : non applicable

Clusters Pig en mode interactif ou en mode de traitement par lots

Amazon EMR vous permet d'exécuter des scripts Pig dans deux modes :

- Interactive
- Par lots

Lorsque vous lancez un cluster de longue durée à l'aide de la console ou de l', vous pouvez vous connecter à l'aide de la console ou de l' AWS CLI, vous pouvez vous connecter à l'aide ssh de la console ou de l', vous pouvez vous connecter à l'aide de la console ou de l', vous pouvez vous connecter à l'aide de la ssu de l', de l', de l', de l', de l', Utiliser Pig de façon interactive vous permet de réviser plus facilement le script Pig qu'en mode de traitement par lots. Lorsque vous avez révisé le script Pig en mode interactif et que tout est correct, vous pouvez charger le script sur Amazon S3 et utiliser le mode de traitement par lots pour exécuter le script en production. Vous pouvez également soumettre des commandes Pig de façon interactive sur un cluster en cours d'exécution pour analyser et transformer les données en fonction des besoins.

En mode de traitement par lots, vous chargez votre script Pig sur Amazon S3, puis soumettez le travail au cluster dans le cadre d'une étape. Les étapes Pig peuvent être soumises à un cluster de longue durée ou à un cluster transitoire.

Caractéristiques de l'application Spark pour les versions AMI antérieures d'Amazon EMR

Utilisation interactive de Spark ou en mode de traitement par lots

Amazon EMR vous permet d'exécuter des applications Spark dans deux modes :

- Interactive
- Par lots

Lorsque vous lancez un cluster de longue durée à l'aide de la console ou de l' AWS CLI, vous pouvez vous connecter à l'aide de la ssH dans le nœud principal en tant qu'utilisateur Hadoop et utiliser le shell Spark pour développer et exécuter vos applications Spark de façon interactive. L'utilisation de Spark vous permet de façon interactive de créer un prototype ou de tester les applications Spark plus facilement que dans un environnement de traitement par lots. Après avoir révisé avec succès l'application Spark en mode interactif, vous pouvez placer le programme JAR ou Python de cette application dans le système de fichiers local du nœud principal du cluster sur Amazon S3. Vous pouvez alors soumettre l'application en tant que flux de travail de traitement par lots.

En mode de traitement par lots, chargez votre script Spark sur Amazon S3 ou le système de fichiers local du nœud principal, puis soumettez le travail au cluster en tant qu'étape. Les étapes Spark peuvent être soumises à un cluster de longue durée ou à un cluster transitoire.

Création d'un cluster avec Spark installé

Pour lancer un cluster avec Spark installé à l'aide de la console

1. Accédez à la nouvelle console Amazon EMR et sélectionnez Changer pour l'ancienne console depuis le menu latéral. Pour plus d'informations sur ce qu'implique le passage à l'ancienne console, consultez la rubrique Utilisation de l'ancienne console.
2. Choisissez Créer un cluster.
3. Pour Configuration des logiciels, choisissez la version AMI dont vous avez besoin.
4. Pour Applications to be installed (Applications à installer), choisissez Spark dans la liste, puis choisissez Configurer et ajouter.
5. Ajoutez des arguments pour modifier la configuration de Spark comme vous le souhaitez. Pour de plus amples informations, veuillez consulter Configuration de Spark. Choisissez Ajouter.
6. Sélectionnez d'autres options si nécessaire, puis choisissez Create cluster (Créer le cluster).

L'exemple suivant montre comment créer un cluster avec Spark au moyen de Java :

```
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
SupportedProductConfig sparkConfig = new SupportedProductConfig()
   .withName("Spark");
```

```
RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("Spark Cluster")
   .withAmiVersion("3.11.0")
   .withNewSupportedProducts(sparkConfig)
   .withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myKeyName")
    .withInstanceCount(1)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m3.xlarge")
    .withSlaveInstanceType("m3.xlarge")
   );
RunJobFlowResult result = emr.runJobFlow(request);
```

Configuration de Spark

Vous configurez Spark lorsque vous créez un cluster en exécutant l'action bootstrap située dans le awslabs/emr-bootstrap-actions/sparkréférentiel sur Github. Pour savoir quels arguments sont acceptés par l'action d'amorçage, consultez le fichier READMEdans ce référentiel. L'action d'amorçage configure les propriétés dans le fichier `$SPARK_CONF_DIR/spark-defaults.conf`. Pour plus d'informations sur les paramètres, consultez la rubrique Spark Configuration dans la documentation Spark. Vous pouvez remplacer « latest » dans l'URL suivante avec le numéro de la version de Spark que vous installez, par exemple, `2.2.0` http://spark.apache.org/docs/latest/configuration.html.

Vous pouvez également configurer Spark dynamiquement au moment de l'envoi de chaque application. Un paramètre permettant d'optimiser automatiquement l'allocation de ressources pour un exécuteur est disponible dans le fichier de configuration `spark`. Pour de plus amples informations, veuillez consulter Remplacement des paramètres de configuration par défaut de Spark.

Modification des paramètres Spark par défaut

L'exemple suivant montre comment créer un cluster avec `spark.executor.memory` défini sur 2G au moyen de l' AWS CLI.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Spark cluster" --ami-version 3.11.0 \
--applications Name=Spark, Args=[-d,spark.executor.memory=2G] --ec2-attributes
 KeyName=myKey \
--instance-type m3.xlarge --instance-count 3 --use-default-roles
```

Soumission d'un travail à Spark

Pour soumettre du travail à un cluster, utilisez une étape pour exécuter le script `spark-submit` sur votre cluster EMR. Ajoutez l'étape en utilisant la `addJobFlowSteps` méthode dans AmazonElasticMapReduceClient:

```
AWSCredentials credentials = new BasicAWSCredentials(accessKey, secretKey);
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
StepFactory stepFactory = new StepFactory();
AddJobFlowStepsRequest req = new AddJobFlowStepsRequest();
req.withJobFlowId("j-1K48XXXXXXHCB");

List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

StepConfig sparkStep = new StepConfig()
 .withName("Spark Step")
 .withActionOnFailure("CONTINUE")
 .withHadoopJarStep(stepFactory.newScriptRunnerStep("/home/hadoop/spark/bin/spark-
submit","--class","org.apache.spark.examples.SparkPi","/home/hadoop/spark/lib/spark-
examples-1.3.1-hadoop2.4.0.jar","10"));

stepConfigs.add(sparkStep);
req.withSteps(stepConfigs);
AddJobFlowStepsResult result = emr.addJobFlowSteps(req);
```

Remplacement des paramètres de configuration par défaut de Spark

Vous souhaiterez probablement remplacer les valeurs de configuration par défaut de Spark indépendamment pour chaque application. Vous pouvez faire cela lorsque vous soumettez des applications à l'aide d'une étape, ce qui transmet essentiellement des options à `spark-submit`. Par exemple, vous pouvez changer la mémoire allouée à un processus d'exécuteur en modifiant `spark.executor.memory`. Vous pouvez fournir le commutateur `--executor-memory` avec un argument tel que :

```
/home/hadoop/spark/bin/spark-submit --executor-memory 1g --class
 org.apache.spark.examples.SparkPi /home/hadoop/spark/lib/spark-examples*.jar 10
```

De même, vous pouvez ajuster `--executor-cores` et `--driver-memory`. Dans une étape, vous pouvez fournir les arguments suivants à l'étape :

```
--executor-memory 1g --class org.apache.spark.examples.SparkPi /home/hadoop/spark/lib/
spark-examples*.jar 10
```

Vous pouvez également ajuster les paramètres qui n'ont peut-être pas de commutateur intégré à l'aide de l'option `--conf`. Pour plus d'informations sur les autres paramètres ajustables, consultez la rubrique Chargement dynamique de propriétés Spark dans la documentation Apache Spark.

DistCp L'utilitaire S3 se différencie des versions AMI antérieures d'Amazon EMR

DistCp Versions de S3 prises en charge dans Amazon EMR

Les DistCp versions S3 suivantes sont prises en charge dans les versions d'AMI Amazon EMR. DistCpLes versions S3 postérieures à 1.0.7 se trouvent directement sur les clusters. Utilisez le fichier JAR dans `/home/hadoop/lib` pour bénéficier des dernières fonctions.

Version	Description	Date de publication
1.0.8	Ajoute les options `--appendToLastFile` , `--requirePreviousManifest` et `--storageClass` .	3 janvier 2014
1.0.7	Ajoute l'option `--s3ServerSideEncryption` .	2 mai 2013
1.0.6	Ajoute l'option `--s3Endpoint` .	6 août 2012
1.0.5	Améliore la capacité à spécifier la version de S3 DistCp à exécuter.	27 juin 2012
1.0.4	Améliore l'option `--deleteOnSuccess` .	19 juin 2012
1.0.3	Ajoute une prise en charge pour les options `--numberFiles` et `--startingIndex` .	12 juin 2012
1.0.2	Améliore l'affectation des noms lors de l'utilisation de groupes.	6 juin 2012

Version	Description	Date de publication
1.0.1	Version initiale de S3DistCp.	19 janvier 2012

Ajout d'une étape de DistCp copie S3 à un cluster

Pour ajouter une étape de DistCp copie S3 à un cluster en cours d'exécution, tapez la commande suivante, remplacez *j-3GYXXXXXX9IOK* par votre ID de cluster et remplacez *amzn-s3-demo-bucket* par votre nom de compartiment Amazon S3.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr add-steps --cluster-id j-3GYXXXXXX9IOK \
--steps Type=CUSTOM_JAR,Name="S3DistCp step",Jar=/home/hadoop/lib/emr-s3distcp-1.0.jar,
\
Args=["--s3Endpoint,s3-eu-west-1.amazonaws.com",\
"--src,s3://amzn-s3-demo-bucket/logs/j-3GYXXXXXX9IOJ/node/",\
"--dest,hdfs:///output",\
"--srcPattern,.*[a-zA-Z,]+"]
```

Example Charger CloudFront les journaux Amazon dans HDFS

Cet exemple charge les CloudFront journaux Amazon dans HDFS en ajoutant une étape à un cluster en cours d'exécution. Au cours du processus, il modifie le format de compression de Gzip (format CloudFront par défaut) en LZO. Cela est utile, car les données compressées à l'aide de LZO peuvent être divisées en plusieurs mappages lorsqu'elles sont décompressées, si bien que vous n'avez pas à attendre la fin de la compression, comme vous le faites pour Gzip. Cela fournit de meilleures performances lorsque vous analysez les données à l'aide d'Amazon EMR. Cet exemple améliore également les performances en utilisant l'expression régulière spécifiée dans l'option `--groupBy` pour combiner tous les journaux d'une heure donnée dans un seul fichier. Les clusters Amazon EMR sont plus efficaces lors du traitement de quelques fichiers volumineux compressés LZO que lors du

traitement de nombreux petits fichiers compressés au format GZIP. Pour fractionner les fichiers LZO, vous devez les indexer et utiliser la bibliothèque tierce hadoop-lzo.

Pour charger CloudFront les journaux Amazon dans HDFS, tapez la commande suivante, remplacez *j-3GYXXXXXX9IOK* par votre ID de cluster et remplacez *amzn-s3-demo-bucket* par votre nom de compartiment Amazon S3.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr add-steps --cluster-id j-3GYXXXXXX9IOK \
--steps Type=CUSTOM_JAR,Name="S3DistCp step",Jar=/home/hadoop/lib/emr-s3distcp-1.0.jar,
\
Args=["--src,s3://amzn-s3-demo-bucket/cf","--dest,hdfs:///local",\
"--groupBy,.*XABCD12345678.([0-9]+-[0-9]+-[0-9]+-[0-9]+).*",\
"--targetSize,128",
"--outputCodec,lzo","--deleteOnSuccess"]
```

Considérons que l'exemple précédent est exécuté sur les fichiers journaux CloudFront suivants.

```
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-01.HLUS3JKx.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-01.I9CNAZrg.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-02.YRRwERSA.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-02.dshVLXFE.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-02.LpLfuShd.gz
```

S3 DistCp copie, concatène et compresse les fichiers dans les deux fichiers suivants, le nom du fichier étant déterminé par la correspondance établie par l'expression régulière.

```
hdfs:///local/2012-02-23-01.lzo
hdfs:///local/2012-02-23-02.lzo
```

Nouveautés

Cette page décrit les modifications et les fonctionnalités disponibles dans les dernières versions d'Amazon EMR 7.x, 6.x et 5.x.

Ces notes de mise à jour sont également disponibles sur les pages Amazon EMR 7.9.0, Amazon EMR 6.15.0 et Amazon EMR 5.36.2, ainsi que les versions de l'application, les versions des composants et les classifications de configuration disponibles pour chaque version.

- Pour les notes de publication des versions précédentes, consultez Archive des notes de publication Amazon EMR.
- Abonnez-vous au flux RSS des notes de mise à jour d'Amazon EMR pour recevoir des mises à jour lorsqu'une nouvelle version d'Amazon EMR est disponible.

> **Note**
>
> Les versions ultérieures d'Amazon EMR utilisent AWS Signature Version 4 (SigV4) pour authentifier les demandes adressées à Amazon S3. Nous vous recommandons d'utiliser une version Amazon EMR compatible avec SigV4 afin de pouvoir accéder aux nouveaux compartiments S3 et d'éviter toute interruption de vos charges de travail. Pour plus d'informations et accéder à la liste des versions d'Amazon EMR compatibles avec SigV4, voir la rubrique Amazon EMR et AWS Signature, version 4.

Amazon EMR 7.9.0 (dernière version de la série 7.x)

Les nouvelles versions d'Amazon EMR sont mises à disposition dans différentes régions sur une période de plusieurs jours, en commençant par la première région à la date de sortie initiale. Il est possible que la dernière version ne soit pas disponible dans votre région pendant cette période.

Les notes de mise à jour suivantes incluent des informations relatives à la version 7.9.0 d'Amazon EMR.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 7.9.0 application upgrades include AWS SDK for Java 2.31.16/1.12.782, Delta 3.3.0-amzn-1, Hadoop 3.4.1, Hudi 0.15.0-amzn-6, Iceberg 1.7.1-amzn-2, Spark 3.5.5, and Zookeeper 3.9.3.
- Support CM-CMK pour l'interface utilisateur persistante de l'application : cette fonctionnalité étend le chiffrement SSE appliqué aux journaux gérés dans le compartiment de service EMR afin de respecter la clé principale personnalisée gérée par le client (CM-CMK). Cette fonctionnalité permet aux clients de gérer l'accessibilité des journaux après la fin de leurs clusters.

Problèmes connus

- Lorsque l'option de journalisation pour EMR activé EC2 est désactivée, les journaux ne sont pas disponibles après la fermeture du cluster et l'application persistante UIs ne contient aucun journal.

Modifications, améliorations et problèmes résolus

- Obsolète d'Oozie — OSS Oozie étant passé en mode attique, la dernière version à inclure Apache Oozie est la version 7.9 d'Amazon EMR. Apache Oozie sera supprimé des versions 7.10 et supérieures d'Amazon EMR. Nous vous recommandons d'utiliser MWAA comme alternative à Oozie. Les clients peuvent se référer à l'utilitaire pour GitHub migrer oozie-to-airflow-emrles flux de travail Oozie existants vers MWAA.
- Emballage de conteneurs YARN — À partir de la version 7.9.0 d'Amazon EMR, la politique d'emballage de conteneurs est désormais disponible pour le planificateur de capacité YARN, qui repose sur la politique de placement de nœuds multiples de YARN.
- Le script hudi-cli.sh — Le script hudi-cli.sh est devenu obsolète dans les versions 7.9.0 et supérieures d'EMR. Les versions 7.9.0 et supérieures d'Amazon EMR utilisent le fichier .jar. hudi-cli-bundle
- Clusters EMR sur l'AMI AL2 023 compatible FIPS — À partir de la version 7.9 d'EMR, les clients peuvent lancer des clusters EMR sur l'AMI 023 compatible FIPS. AL2 Cette modification met à jour les signatures numériques et les algorithmes de hachage d'EMR afin de garantir que tous les packages EMR respectent les politiques cryptographiques recommandées pour la norme FIPS 140-3. Remarque : Cela ne couvre pas le statut de certification FIPS EMR AL2 ou 023. Reportez-vous AL2à la FAQ FIPS 023 pour obtenir les dernières informations sur le statut de validation de la norme FIPS 140-3.

- Latence par étapes : cette version résout un problème selon lequel la latence des étapes augmentait lors de l'utilisation du chiffrement en transit sans sélectionner explicitement l'application Hadoop à installer.
- Lancement de clusters avec des volumes chiffrés EBS : cette version résout un problème selon lequel un appel KMS sur des volumes chiffrés par EBS pouvait entraîner l'échec du lancement d'un cluster EMR.

-

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023.7.2 250527.1	6.1.134-152.225 mai 2023	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californ ie du Nord), Asie-Paci fique (Singapour), Asie-Pacifique (Sydney), Israël (Tel Aviv), Canada (centre), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,7.2 250331.(	6.1.131-143.221 mai 2023	18 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre), Asie-Pacifique (Tokyo)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250303.(	6.1.129-138.220 août 2023	27 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250218,:	6.1.128-136.201. mai 2023	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2023,6.2 250211.(	6.1.127-135.201. mai 2023	26 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary), Asie-Pacifique (Malaisie), Asie-Paci fique (Malaisie), Mexique (centre)

Amazon EMR 6.15.0 (dernière version de la série 6.x)

Les nouvelles versions d'Amazon EMR sont mises à disposition dans différentes régions sur une période de plusieurs jours, en commençant par la première région à la date de sortie initiale. Il est possible que la dernière version ne soit pas disponible dans votre région pendant cette période.

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.15.0. Les modifications ont été apportées à la version 6.14.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.15.0.

Nouvelles fonctionnalités

- Mises à niveau des applications –Amazon EMR 6.15.0 application upgrades include Apache Hadoop 3.3.6, Apache Hudi 0.14.0-amzn-0, Iceberg 1.4.0-amzn-0, and Trino 426..
- Lancement plus rapide pour les clusters EMR exécutés sur EC2 — Le lancement d'un cluster Amazon EMR sur un cluster est désormais jusqu'à 35 % plus rapide. EC2 Grâce à cette amélioration, la plupart des clients peuvent lancer leurs clusters en moins de 5 minutes.
- CodeWhisperer pour EMR Studio — Vous pouvez désormais utiliser Amazon avec CodeWhisperer Amazon EMR Studio pour obtenir des recommandations en temps réel lorsque vous écrivez du code. JupyterLab CodeWhisperer peut compléter vos commentaires, terminer des lignes de code individuelles, faire line-by-line des recommandations et générer des fonctions entièrement formées.
- Redémarrage accéléré des tâches avec Flink – Avec Amazon EMR 6.15.0 et versions ultérieures, plusieurs nouveaux mécanismes sont disponibles pour Apache Flink afin d'améliorer le temps de

redémarrage des tâches lors des opérations de récupération ou de mise à l'échelle des tâches. Cela permet d'optimiser la vitesse de récupération et de redémarrage des graphes d'exécution afin d'améliorer la stabilité des tâches.

- Contrôle d'accès détaillé et au niveau des tables pour les formats de tables ouvertes : avec Amazon EMR 6.15.0 et versions ultérieures, lorsque vous exécutez des tâches Spark sur Amazon EMR sur des clusters EC2 qui accèdent aux données du catalogue de données AWS Glue, vous pouvez appliquer des autorisations au niveau des tables, des lignes, des colonnes et des cellules AWS Lake Formation aux tables basées sur Hudi, Iceberg ou Delta Lake.
- Mise à niveau de Hadoop – Amazon EMR 6.15.0 inclut une mise à niveau d'Apache Hadoop vers la version 3.3.6. Hadoop 3.3.6 était la dernière version au moment du déploiement d'Amazon EMR 6.15, publiée par Apache en juin 2023. Les versions précédentes d'Amazon EMR (6.9.0 à 6.14.x) utilisaient Hadoop 3.3.3.

  La mise à niveau inclut des centaines d'améliorations et de correctifs, ainsi que des fonctionnalités telles que des paramètres de nœud de données reconfigurables, l'option `DFSAdmin` qui permet de lancer des opérations de reconfiguration en bloc sur tous les nœuds de données actifs et une API vectorielle qui permet aux lecteurs à forte densité de recherche de spécifier plusieurs plages à lire. Hadoop 3.3.6 ajoute également la prise en charge du HDFS APIs et de la sémantique à son journal d'écriture anticipée (WAL), afin qu' HBase il puisse fonctionner sur d'autres implémentations de systèmes de stockage. Pour plus d'informations, consultez les journaux des modifications pour les versions 3.3.4, 3.3.5 et 3.3.6 dans la documentation Apache Hadoop.
- Support du AWS SDK pour Java, version 2 - Les applications Amazon EMR 6.15.0 peuvent utiliser les versions 1.12.569 ou 2.20.160 du SDK pour AWS Java si l'application prend en charge la version v2. Le AWS SDK pour Java 2.x est une réécriture majeure de la base de code de la version 1.x. Il repose sur Java 8+ et ajoute plusieurs fonctionnalités fréquemment demandées. Cela inclut la prise en charge d'E/S non bloquantes et la possibilité de brancher une implémentation HTTP différente au moment de l'exécution. Pour plus d'informations, notamment un Guide de migration du kit SDK pour la version 1 de Java version la version 2, voir le guide du Kit AWS SDK pour Java, version 2.

Modifications, améliorations et problèmes résolus

- Pour améliorer vos clusters EMR haute disponibilité, cette version assure la connectivité aux démons Amazon EMR sur un hôte local utilisant des points de terminaison IPv6.
- Cette version permet au protocole TLS 1.2 de communiquer avec les nœuds ZooKeeper provisionnés sur tous les nœuds principaux de votre cluster à haute disponibilité.

- Cette version améliore la gestion des fichiers journaux de ZooKeeper transactions conservés sur les nœuds principaux afin de minimiser les scénarios dans lesquels les fichiers journaux dépassent les limites et interrompent les opérations du cluster.
- Cette version accentue la résilience de la communication intra-nœud pour les clusters EMR haute disponibilité. Cette amélioration réduit le risque d'échec des actions d'amorçage ou de démarrage du cluster.
- Tez dans Amazon EMR 6.15.0 introduit des configurations que vous pouvez spécifier pour ouvrir de manière asynchrone les divisions d'entrée dans une division groupée Tez. Cela permet d'accélérer les performances des requêtes de lecture lorsqu'il existe un grand nombre de divisions d'entrée dans une seule division groupée Tez. Pour plus d'informations, consultez la rubrique Ouverture fractionnée asynchrone Tez.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique (Taipei), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Malaisie), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Calgary), Europe (Zurich), Asie-Pacifique (Malaisie)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	13 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

Amazon EMR 5.36.2 (dernière version de la série 5.x)

Les nouvelles versions d'Amazon EMR sont mises à disposition dans différentes régions sur une période de plusieurs jours, en commençant par la première région à la date de sortie initiale. Il est possible que la dernière version ne soit pas disponible dans votre région pendant cette période.

Les notes de mise à jour suivantes incluent des informations relatives à la version 5.36.2 d'Amazon EMR. Les modifications sont relatives à la version 5.36.1. Pour plus d'informations sur le calendrier de publication, consultez le journal des modifications.

Modifications, améliorations et problèmes résolus

- Cette version améliore la logique de réduction du cluster afin qu'Amazon EMR ne réduise pas les nœuds principaux en dessous du paramètre de facteur de réplication HDFS pour le cluster. Cette amélioration répond aux exigences de redondance des données et réduit le risque de blocage d'une opération de dimensionnement.
- Cette version ajoute un nouveau mécanisme de nouvelle tentative au flux de travail de dimensionnement des clusters pour les clusters EMR qui exécutent Presto ou Trino. Cette amélioration réduit le risque que le redimensionnement du cluster s'exécute indéfiniment en raison de l'échec d'une seule opération de redimensionnement. Cela améliore également l'utilisation du cluster, car celui-ci augmente et diminue la capacité plus rapidement.
- Résout un problème selon lequel les opérations de réduction du cluster pouvaient être bloquées alors qu'Amazon EMR mettait hors service gracieusement un nœud principal et que celui-ci devenait inutilisable avant sa mise hors service complète.
- Améliore la stabilité d'un nœud dans un cluster à haute disponibilité comportant plusieurs nœuds principaux lorsqu'Amazon EMR redémarre un seul nœud.

- Optimise la gestion des journaux avec Amazon EMR exécuté sur Amazon. EC2 C'est pourquoi vous constaterez peut-être une légère réduction des coûts de stockage pour les journaux de votre cluster.
- Améliore la gestion des fichiers journaux de ZooKeeper transactions conservés sur les nœuds principaux afin de minimiser les scénarios dans lesquels les fichiers journaux dépassent les limites et interrompent les opérations du cluster.
- Corrige un bogue rare qui peut provoquer l'échec d'un cluster à haute disponibilité comportant plusieurs nœuds principaux en raison de l'impossibilité de communiquer avec le Yarn ResourceManager.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique (Taipei), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Malaisie), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)
2,0.2024 503.0	4,14,343	xxxxxx, 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

Amazon EMR et AWS Signature, version 4

Les versions Amazon EMR utilisent AWS Signature Version 4 (SigV4) pour authentifier les demandes adressées à Amazon S3. Les compartiments créés dans Amazon S3 après le 24 juin 2020 ne prennent pas en charge les demandes signées par Signature Version 2 (SigV2). Les buckets créés le 24 juin 2020 ou avant cette date continueront de prendre en charge le protocole SIGv2. Nous vous recommandons de migrer vers une version Amazon EMR qui prend en charge SigV4 afin de pouvoir accéder aux nouveaux compartiments S3 et d'éviter toute interruption de vos charges de travail.

Si vous utilisez des applications incluses dans Amazon EMR, telles qu'Apache Spark, Apache Hive et Presto, vous n'avez pas besoin de modifier le code de votre application pour utiliser SigV4. Si vous utilisez des applications personnalisées qui ne sont pas incluses dans Amazon EMR, vous devrez peut-être mettre à jour votre code pour utiliser SigV4. Pour plus d'informations, consultez la section Passage de Signature Version 2 à Signature Version 4 dans le Guide de l'utilisateur Amazon S3.

SigV4 est pris en charge par les versions suivantes d'Amazon EMR : emr-4.7.4, emr-4.8.5, emr-4.9.6, emr-4.10.1, emr-5.1.1, emr-5.2.3, emr-5.3.2, emr-5.4.1, emr-5.5.4, emr-5.6.1, emr-5.7.1, emr-5.8.3, emr-5.9.1, emr-5.10.1, emr-5.11.4, emr-5.12.3, emr-5.13.1, emr-5.14.2, emr-5.15.1, emr-5.16.1, emr-5.17.2, emr-5.18.1, emr-5.19.1, emr-5.20.1, emr-5.21.2, and emr-5.22.0 and higher. Toutes les versions 6.x et 7.x prennent en charge SigV4.

Approche visant à atténuer le CVE-2021-44228

> **Note**
>
> Pour Amazon EMR version 6.9.0 et versions ultérieures, tous les composants installés par Amazon EMR qui utilisent les bibliothèques Log4j utilisent Log4j version 2.17.1 ou ultérieure.

Amazon EMR s'exécute sur EC2

Le problème abordé dans CVE-2021-44228 concerne les versions principales d'Apache Log4j comprises entre 2.0.0 et 2.14.1 lors du traitement d'entrées provenant de sources non fiables. Les clusters Amazon EMR lancés avec les versions 5.x d'Amazon EMR 5.x jusqu'à 5.34.0 et les versions d'EMR 6.x antérieures à Amazon EMR 6.5.0 incluent des frameworks open source tels qu'Apache Hive, Flink, HUDI, Presto et Trino, qui utilisent ces versions d'Apache Log4j. Cependant, de nombreux clients utilisent les frameworks open source installés sur leurs clusters Amazon EMR pour traiter et enregistrer les entrées provenant de sources non fiables.

Nous vous recommandons d'appliquer la « solution d'action Amazon EMR Bootstrap pour Log4j CVE-2021-44228 » comme décrit dans la section suivante. Cette solution prend également en charge le CVE-2021-45046.

> **Note**
>
> Les scripts d'action d'amorçage pour Amazon EMR ont été mis à jour le 7 septembre 2022 pour inclure des corrections de bogues progressives et des améliorations pour Oozie. Si vous utilisez Oozie, vous devez appliquer la solution d'action d'amorçage Amazon EMR mise à jour décrite dans la section suivante.

Amazon EMR on EKS

Si vous utilisez Amazon EMR sur EKS avec une configuration par défaut, vous n'êtes pas concerné par le problème décrit dans le CVE-2021-44228 et vous n'êtes pas obligé d'appliquer la solution décrite dans la section Solution d'action d'amorçage Amazon EMR pour Log4j CVE-2021-44228 et CVE-2021-45046. Pour Amazon EMR sur EKS, le moteur d'exécution Amazon EMR pour Spark utilise Apache Log4j version 1.2.17. Lorsque vous utilisez Amazon EMR sur EKS, vous ne devez pas modifier le paramètre par défaut du composant `log4j.appender` sur `log`.

Solution d'action d'amorçage Amazon EMR pour Log4j CVE-2021-44228 et CVE-2021-45046

Cette solution fournit une action d'amorçage Amazon EMR qui doit être appliquée à vos clusters Amazon EMR. Pour chaque version d'Amazon EMR, vous trouverez ci-dessous un lien vers un script d'action d'amorçage. Pour appliquer cette action d'amorçage, vous devez suivre les étapes suivantes :

1. Copiez le script correspondant à votre version d'Amazon EMR dans un compartiment S3 local de votre Compte AWS. Assurez-vous que vous utilisez un script d'amorçage spécifique à votre version d'Amazon EMR.
2. Configurez une action d'amorçage pour vos clusters EMR afin d'exécuter le script copié dans votre compartiment S3 conformément aux instructions décrites dans Documentation EMR. Si d'autres actions d'amorçage sont configurées pour vos clusters EMR, assurez-vous que ce script est configuré comme le premier script d'action d'amorçage à exécuter.
3. Mettez fin aux clusters EMR existants et lancez de nouveaux clusters à l'aide du script d'action bootstrap. AWS recommande de tester les scripts bootstrap dans votre environnement de test et

de valider vos applications avant de les appliquer à votre environnement de production. Si vous n'utilisez pas la dernière version d'une version mineure d'EMR (par exemple, 6.3.0), vous devez utiliser la dernière révision (par exemple, 6.3.1), puis appliquer la solution décrite ci-dessus.

CVE-2021-44228 et CVE-2021-45046 – Scripts d’amorçage pour les versions d’Amazon EMR

Numéro de version d’Amazon EMR	Emplacement du script	Date de sortie du script
6.5.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.5.0-v2.sh`	24 mars 2022
6.4.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.4.0-v2.sh`	24 mars 2022
6.3.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.3.1-v2.sh`	24 mars 2022
6.2.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.2.1-v2.sh`	24 mars 2022
6.1.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.1.1-v2.sh`	14 décembre 2021
6.0.1	`s3://elasticmapreduce/bootstrap-actions/`	14 décembre 2021

Numéro de version d'Amazon EMR	Emplacement du script	Date de sortie du script
	`log4j/patch-log4j-emr-6.0.1-v2.sh`	
5,34,0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.34.0-v2.sh`	12 décembre 2021
5,33.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.33.1-v2.sh`	12 décembre 2021
5,32.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.32.1-v2.sh`	13 décembre 2021
5,31.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.31.1-v2.sh`	13 décembre 2021
5,30,2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.30.2-v2.sh`	14 décembre 2021
5,29,0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.29.0-v2.sh`	14 décembre 2021

Numéro de version d'Amazon EMR	Emplacement du script	Date de sortie du script
5,28.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.28.1-v2.sh`	15 décembre 2021
5,27.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.27.1-v2.sh`	15 décembre 2021
5,26,0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.26.0-v2.sh`	15 décembre 2021
5,25,0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.25.0-v2.sh`	15 décembre 2021
5.24.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.24.1-v2.sh`	15 décembre 2021
5.23.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.23.1-v2.sh`	15 décembre 2021

Numéro de version d'Amazon EMR	Emplacement du script	Date de sortie du script
5,22,0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.22.0-v2.sh`	15 décembre 2021
5.21.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.21.2-v2.sh`	15 décembre 2021
5.20.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.20.1-v2.sh`	15 décembre 2021
5.19.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.19.1-v2.sh`	15 décembre 2021
5.18.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.18.1-v2.sh`	15 décembre 2021
5,17.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.17.2-v2.sh`	15 décembre 2021

Numéro de version d'Amazon EMR	Emplacement du script	Date de sortie du script
5.16.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.16.1-v2.sh`	15 décembre 2021
5.15.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.15.1-v2.sh`	15 décembre 2021
5.14.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.14.2-v2.sh`	15 décembre 2021
5.13.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.13.1-v2.sh`	15 décembre 2021
5.12.3	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.12.3-v2.sh`	15 décembre 2021
5.11.4	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.11.4-v2.sh`	15 décembre 2021

Numéro de version d'Amazon EMR	Emplacement du script	Date de sortie du script
5.10.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.10.1-v2.sh`	15 décembre 2021
5.9.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.9.1-v2.sh`	15 décembre 2021
5.8.3	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.8.3-v2.sh`	15 décembre 2021
5.7.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.7.1-v2.sh`	15 décembre 2021

Version EMR	Dernière révision en date de décembre 2021
6.3.0	6.3.1
6.2.0	6.2.1
6,10	6.1.1
6.0.0	6.0.1
5,33,0	5,33.1
5,32,0	5,32.1

Version EMR	Dernière révision en date de décembre 2021
5,31,0	5,31.1
5.30.0 ou 5.30.1	5,30,2
5,28,0	5,28.1
5,27,0	5,27.1
5,24,0	5.24.1
5.23.0	5.23.1
5.21.0 ou 5.21.1	5.21.2
5,20,0	5.20.1
5,19,0	5.19.1
5,18,0	5.18.1
5.17.0 ou 5.17.1	5,17.2
5.16.0	5.16.1
5.15.0	5.15.1
5.14.0 ou 5.14.1	5.14.2
5.13.0	5.13.1
5.12.0, 5.12.1, 5.12.2	5.12.3
5.11.0, 5.11.1, 5.11.2, 5.11.3	5.11.4
5.9.0	5.9.1
5,8.0, 5,8.1, 5,8.2	5.8.3
5.7.0	5.7.1

Questions fréquentes (FAQ)

- Les versions d'EMR antérieures à EMR 5 sont-elles affectées par le CVE-2021-44228 ?

  Non Les versions EMR antérieures à la version 5 d'EMR utilisent des versions de Log4j antérieures à 2.0.

- Cette solution répond-elle au CVE-2021-45046 ?

  Oui, cette solution prend également en charge le CVE-2021-45046.

- La solution gère-t-elle les applications personnalisées que j'installe sur mes clusters EMR ?

  Le script d'amorçage met uniquement à jour les fichiers JAR installés par EMR. Si vous installez et exécutez des applications personnalisées et des fichiers JAR sur vos clusters EMR par le biais d'actions d'amorçage, d'étapes soumises à vos clusters, en utilisant une AMI Amazon Linux personnalisée ou via tout autre mécanisme, contactez votre fournisseur d'applications pour déterminer si vos applications personnalisées sont affectées par le CVE-2021-44228 et pour déterminer une solution appropriée.

- Comment gérer les images docker personnalisées avec EMR sur EKS ?

  Si vous ajoutez des applications personnalisées à Amazon EMR on EKS à l'aide d'images docker personnalisées ou si vous soumettez des tâches à Amazon EMR sur des fichiers d'application EKSwith personnalisés, contactez le fournisseur de l'application pour déterminer si vos applications personnalisées sont concernées par le CVE-2021-44228 et pour déterminer une solution appropriée.

- Comment fonctionne le script bootstrap pour atténuer le problème décrit dans CVE-2021-44228 et CVE-2021-45046 ?

  Le script d'amorçage met à jour les instructions de démarrage d'EMR en ajoutant un nouveau jeu d'instructions. Ces nouvelles instructions suppriment les fichiers JndiLookup de classe utilisés via Log4j par tous les frameworks open source installés par EMR. Cela fait suite à la recommandation publiée par Apache pour résoudre les problèmes liés à Log4j.

- Existe-t-il une mise à jour d'EMR qui utilise les versions 2.17.1 ou supérieures de Log4j ?

  Les versions EMR 5 jusqu'à la version 5.34 et les versions EMR 6 jusqu'à la version 6.5 utilisent d'anciennes versions de frameworks open source incompatibles avec les dernières versions de Log4j. Si vous continuez à utiliser ces versions, nous vous recommandons d'appliquer l'action bootstrap pour atténuer les problèmes décrits dans le CVEs. Après EMR 5 version 5.34 et EMR 6

version 6.5, les applications qui utilisent Log4j 1.x et Log4j 2.x seront mises à niveau pour utiliser Log4j 1.2.17 (ou supérieur) et Log4j 2.17.1 (ou supérieur) respectivement, et ne nécessiteront pas l'utilisation des actions d'amorçage fournies ci-dessus pour atténuer les problèmes CVE.

- Les versions d'EMR sont-elles affectées par le CVE-2021-45105 ?

  Les applications installées par Amazon EMR avec les configurations par défaut d'EMR ne sont pas affectées par CVE-2021-45105. Parmi les applications installées par Amazon EMR, seul Apache Hive utilise Apache Log4j avec des recherches contextuelles, et il n'utilise pas de disposition de modèle autre que celle par défaut de manière à permettre le traitement de données d'entrée inappropriées.

- Amazon EMR est-il concerné par l'une des divulgations CVE suivantes ?

  Le tableau suivant contient une liste de ceux CVEs qui sont liés à Log4j et indique si chaque CVE a un impact sur Amazon EMR. Les informations de ce tableau ne s'appliquent que lorsque les applications sont installées par Amazon EMR à l'aide des configurations par défaut.

CVE	Incidence sur EMR	Remarques
CVE-2022-23302	Non	Amazon EMR ne configure pas Log4j JMSSink
CVE-2022-23305	Non	Amazon EMR ne configure pas Log4j JDBCAppender
CVE-2022-23307	Non	Amazon EMR ne configure pas Log4j Chainsaw
CVE-2020-9493	Non	Amazon EMR ne configure pas Log4j Chainsaw
CVE-2021-44832	Non	Amazon EMR ne configure pas Log4j JDBCAppender avec une chaîne de connexion JNDI
CVE-2021-4104	Non	Amazon EMR n'utilise pas Log4j JMSAppender

CVE	Incidence sur EMR	Remarques
CVE-2020-9488	Non	Les applications installées par Amazon EMR n'utilisent pas Log4j SMTPAppender
CVE-2019-17571	Non	Amazon EMR bloque l'accès public aux clusters et ne se lance pas SocketServer
CVE-2019-17531	Non	Nous vous recommandons de passer à la dernière version Amazon EMR. Amazon EMR 5.33.0 et versions ultérieures utilisent jackson-databind 2.6.7.4 ou version ultérieure, et EMR 6.1.0 et versions ultérieures utilisent jackson-databind 2.10.0 ou version ultérieure. Ces versions de jackson-databind ne sont pas affectées par le CVE.

Archive des notes de publication Amazon EMR

Les notes de mise à jour de toutes les versions d'Amazon EMR sont disponibles ci-dessous. Pour obtenir des informations complètes sur chaque version, consultez Versions Amazon EMR 6.x, Versions Amazon EMR 5.x et Versions Amazon EMR 4.x.

Abonnez-vous au flux RSS des notes de mise à jour d'Amazon EMR pour recevoir des mises à jour lorsqu'une nouvelle version d'Amazon EMR est disponible.

Version 6.14.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.14.0. Les modifications ont été apportées à la version 6.13.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.14.0.

Nouvelles fonctionnalités

- Amazon EMR 6.14.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-2, Flink 1.17.1, Iceberg 1.3.1, and Trino 422.
- La mise à l'échelle gérée par Amazon EMR est désormais disponible dans la région `ap-southeast-3` Asie-Pacifique (Jakarta) pour les clusters que vous créez avec Amazon EMR 6.14.0 ou version ultérieure.

Modifications, améliorations et problèmes résolus

- La version 6.14.0 optimise la gestion des journaux avec Amazon EMR exécuté sur Amazon. EC2 C'est pourquoi vous constaterez peut-être une légère réduction des coûts de stockage pour les journaux de votre cluster.
- La version 6.14.0 améliore le flux de travail de mise à l'échelle afin de prendre en compte les différentes instances principales dont la taille varie considérablement pour leurs volumes Amazon EBS. Cette amélioration s'applique uniquement aux nœuds principaux ; les opérations de réduction des nœuds de tâches ne sont pas affectées.
- La version 6.14.0 améliore la façon dont Amazon EMR interagit avec les applications open source telles que Apache Hadoop YARN ResourceManager and HDFS NameNode. Cette amélioration réduit le risque de retards opérationnels liés à la mise à l'échelle du cluster et atténue les échecs de démarrage dus à des problèmes de connectivité avec les applications open source.
- La version 6.14.0 optimise l'installation des applications au lancement du cluster. Cela améliore les temps de démarrage du cluster pour certaines combinaisons d'applications Amazon EMR.
- La version 6.14.0 corrige un problème où les opérations de réduction de cluster pouvaient se bloquer lorsqu'un cluster exécuté dans un VPC avec un domaine personnalisé rencontrait un redémarrage de nœud principal ou de nœud de tâche.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique (Taipei), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Malaisie), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Calgary), Europe (Zurich), Asie-Pacifique (Malaisie)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	17 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	11 septembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

Version 6.13.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.13.0. Les modifications ont été apportées à la version 6.12.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.13.0.

Nouvelles fonctionnalités

- Amazon EMR 6.13.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-1, CUDA Toolkit 11.8.0, and JupyterHub 1.5.0.

Modifications, améliorations et problèmes résolus

- La version 6.13.0 améliore le démon de gestion des journaux Amazon EMR afin de garantir que tous les journaux sont chargés à une cadence régulière sur Amazon S3 lorsqu'une commande de résiliation de cluster est émise. Cela permet de résilier plus rapidement les clusters.
- La version 6.13.0 améliore les capacités de gestion des journaux Amazon EMR afin de garantir un chargement cohérent et en temps voulu de tous les fichiers journaux sur Amazon S3. Cela est particulièrement avantageux pour les clusters EMR de longue durée.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

Version 6.12.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.12.0. Les modifications ont été apportées à la version 6.11.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.12.0.

Nouvelles fonctionnalités

- Amazon EMR 6.12.0 supports Apache Spark 3.4.0, Apache Spark RAPIDS 23.06.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.1-amzn-0, Apache Iceberg 1.3.0-amzn-0, Trino 414, and PrestoDB 0.281.
- Les versions 6.12.0 et supérieures d'Amazon EMR prennent en charge l'intégration LDAP avec Apache Livy, Apache Hive via HiveServer 2 (HS2), Trino, Presto et Hue. Vous pouvez également installer Apache Spark et Apache Hadoop sur un cluster EMR utilisant la version 6.12.0 ou supérieure et les configurer pour utiliser LDAP. Pour plus d'informations, consultez Utilisation de serveurs Active Directory ou LDAP pour l'authentification avec Amazon EMR.

Modifications, améliorations et problèmes résolus

- Les versions 6.12.0 et supérieures d'Amazon EMR fournissent la prise en charge de l'exécution Java 11 pour Flink. Pour de plus amples informations, veuillez consulter Configurer Flink pour qu'il fonctionne avec Java 11.
- La version 6.12.0 ajoute un nouveau mécanisme de nouvelle tentative au flux de travail de dimensionnement des clusters pour les clusters EMR qui exécutent Presto ou Trino. Cette amélioration réduit le risque que le redimensionnement du cluster soit bloqué indéfiniment en raison de l'échec d'une seule opération de redimensionnement. Cela améliore également l'utilisation du cluster, car celui-ci augmente et diminue la capacité plus rapidement.
- La version 6.12.0 corrige un problème à cause duquel les opérations de réduction de la taille du cluster peuvent se bloquer lorsqu'un nœud principal en cours de mise hors service appropriée tombe en panne pour une raison quelconque avant d'être complètement mis hors service.
- La version 6.12.0 améliore la logique de réduction de la taille des clusters afin que votre cluster ne tente pas une réduction d'échelle des nœuds principaux en dessous du paramètre de facteur de réplication HDFS défini pour le cluster. Cela répond à vos exigences en matière de redondance des données et réduit le risque de blocage d'une opération de dimensionnement.
- La version 6.12.0 améliore les performances et l'efficacité du service de surveillance de l'état d'Amazon EMR en augmentant la vitesse à laquelle il enregistre les changements d'état des

instances. Cette amélioration réduit le risque de dégradation des performances des nœuds de cluster qui exécutent plusieurs outils clients personnalisés ou des applications tierces.

- La version 6.12.0 améliore les performances du démon de gestion des journaux sur le cluster pour Amazon EMR. Par conséquent, les risques de dégradation des performances sont moindres avec les clusters EMR qui exécutent des étapes avec une simultanéité élevée.
- Avec la version 6.12.0 d'Amazon EMR, le démon de gestion des journaux a été mis à niveau pour identifier tous les journaux en cours d'utilisation avec des descripteurs de fichiers ouverts sur le stockage d'instance local, ainsi que les processus associés. Cette mise à niveau garantit qu'Amazon EMR supprime correctement les fichiers et récupère de l'espace de stockage une fois les journaux archivés dans Amazon S3.
- La version 6.12.0 inclut une amélioration du démon de gestion des journaux qui supprime les répertoires d'étapes vides et inutilisés dans le système de fichiers du cluster local. Un trop grand nombre de répertoires vides peut dégrader les performances des démons Amazon EMR et entraîner une surutilisation du disque.
- La version 6.12.0 permet la rotation des journaux pour les journaux YARN Timeline Server. Cela permet de minimiser les scénarios de surutilisation des disques, en particulier pour les clusters de longue durée.
- La taille du volume racine par défaut est passée à 15 Go dans Amazon EMR 6.10.0 et versions ultérieures. Les versions antérieures ont une taille de volume racine par défaut de 10 Go.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 822.0	4,1,4322	30 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre)

Version 6.11.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.11.1. Les modifications ont été apportées à la version 6.11.0. Pour plus d'informations sur le calendrier de publication, consultez le Journal des modifications 6.11.1.

Modifications, améliorations et problèmes résolus

- En raison d'un conflit de verrouillage, un nœud peut se retrouver bloqué s'il est ajouté ou supprimé en même temps qu'il est mis hors service. Par conséquent, le gestionnaire de ressources Hadoop (YARN) ne répond plus et affecte tous les conteneurs entrants et en cours d'exécution.
- Cette version inclut une modification qui permet aux clusters à haute disponibilité de se remettre d'un état défaillant après le redémarrage.
- Cette version inclut des correctifs de sécurité pour Hue et HBase.
- Cette version résout un problème selon lequel les clusters exécutant des charges de travail sur Spark avec Amazon EMR peuvent recevoir silencieusement des résultats incorrects avec `contains`, `startsWith`, `endsWith` et `like`. Ce problème se produit lorsque vous utilisez les expressions sur des champs partitionnés contenant des métadonnées dans Amazon EMR Hive3 Metastore Server (HMS).
- Cette version corrige un problème de limitation du côté de Glue en l'absence de fonctions définies par l'utilisateur (UDF).
- Cette version corrige un problème qui supprime les journaux des conteneurs par le service d'agrégation des journaux des nœuds avant que le transmetteur de journaux ne puisse les envoyer vers S3 en cas de mise hors service de YARN.
- Cette version corrige un problème lié aux métriques du FairShare planificateur lorsque Node Label est activé pour Hadoop.
- Cette version corrige un problème qui affectait les performances de Spark lorsque vous définissez une valeur `true` par défaut pour la configuration `spark.yarn.heterogeneousExecutors.enabled` dans `spark-defaults.conf`.
- Cette version corrige un problème lié à l'échec de la lecture des données de shuffle par Reduce Task. Ce problème provoquait des échecs de requêtes Hive avec une erreur de mémoire corrompue.
- Cette version ajoute un nouveau mécanisme de nouvelle tentative au flux de travail de dimensionnement des clusters pour les clusters EMR qui exécutent Presto ou Trino. Cette amélioration réduit le risque que le redimensionnement du cluster soit bloqué indéfiniment

en raison de l'échec d'une seule opération de redimensionnement. Cela améliore également l'utilisation du cluster, car celui-ci augmente et diminue la capacité plus rapidement.

- Cette version améliore la logique de réduction de la taille des clusters afin que votre cluster ne tente pas une réduction d'échelle des nœuds principaux en dessous du paramètre de facteur de réplication HDFS défini pour le cluster. Cela répond à vos exigences en matière de redondance des données et réduit le risque de blocage d'une opération de dimensionnement.
- Le démon de gestion des journaux a été mis à niveau pour identifier tous les journaux en cours d'utilisation avec des descripteurs de fichiers ouverts sur le stockage d'instance local, ainsi que les processus associés. Cette mise à niveau garantit qu'Amazon EMR supprime correctement les fichiers et récupère de l'espace de stockage une fois les journaux archivés dans Amazon S3.
- Cette version inclut une amélioration du démon de gestion des journaux qui supprime les répertoires d'étapes vides et inutilisés dans le système de fichiers du cluster local. Un trop grand nombre de répertoires vides peut dégrader les performances des démons Amazon EMR et entraîner une surutilisation du disque.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn) , Chine (Pékin), Asie-Pacifique (Malaisie), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), USA Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyde) Derabad), États-Unis Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Malaisie), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis Est), (États-Unis Ouest) AWS GovCloud , États-Unis Ouest (AWS GovCloud Oregon), Mexique (centre), États-Unis Ouest (Californie du Nord), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Asie-Pacifique (Taipei), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Malaisie), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Malaisie), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (centre), Canada

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest (Calgary) AWS GovCloud , (États-Unis est), (États-Unis ouest) AWS GovCloud , Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv), Mexique (centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Calgary), Europe (Zurich), Asie-Pacifique (Malaisie)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA Ouest)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud , (USA Est), Chine (Pékin), Chine (Ningxia), Canada Ouest (Calgary)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Asie-Paci fique (Melbourne), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn) , Moyen-Orient (Émirats arabes unis), Canada (Centre), Israël (Tel Aviv) AWS GovCloud , (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), (USA Est) AWS GovCloud , Chine (Pékin), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 822.0	4,1,4322	30 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

Version 6.11.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.11.0. Les modifications ont été apportées à la version 6.10.0. Pour plus d'informations sur le calendrier de publication, consultez le journal des modifications.

Nouvelles fonctionnalités

- Amazon EMR 6.11.0 prend en charge Apache Spark 3.3.2-amzn-0, Apache Spark RAPIDS 23.02.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.0-amzn-0, Apache Iceberg 1.2.0-amzn-0, Trino 410-amzn-0 et PrestoDB 0.279-amzn-0.

Modifications, améliorations et problèmes résolus

- Avec Amazon EMR 6.11.0, le connecteur DynamoDB a été mis à niveau vers la version 5.0.0. La version 5.0.0 utilise AWS SDK for Java 2.x. Les versions précédentes utilisaient la version AWS SDK pour Java 1.x. À la suite de cette mise à niveau, nous vous conseillons vivement de tester votre code avant d'utiliser le connecteur DynamoDB avec Amazon EMR 6.11.
- Lorsque le connecteur DynamoDB pour Amazon EMR 6.11.0 appelle le service DynamoDB, il utilise la valeur de région que vous fournissez pour la propriété `dynamodb.endpoint`. Nous vous recommandons de configurer également `dynamodb.region` lorsque vous utilisez `dynamodb.endpoint`, et que les deux propriétés ciblent la même Région AWS. Si vous l'utilisez `dynamodb.endpoint` et que vous ne le configurez pas`dynamodb.region`, le connecteur DynamoDB pour Amazon EMR 6.11.0 renverra une exception de région non valide et tentera de réconcilier vos Région AWS informations avec le service de métadonnées d'instance Amazon EC2 (IMDS). Si le connecteur ne parvient pas à récupérer la région depuis IMDS, la valeur par défaut est USA Est (Virginie du Nord) (`us-east-1`). L'erreur suivante illustre l'exception de région non valide que vous pourriez obtenir si vous ne configurez pas correctement la `dynamodb.region` propriété : `error software.amazon.awssdk.services.dynamodb.model.DynamoDbException: Credential should be scoped to a valid region.` pour plus d'informations sur les classes concernées par la AWS SDK pour Java mise à niveau vers la version 2.x, consultez le commit Upgrade AWS SDK pour Java from 1.x to 2.x (#175) dans le GitHub dépôt du connecteur Amazon EMR - DynamoDB.
- Cette version corrige un problème où les données des colonnes deviennent `NULL` lorsque vous utilisez Delta Lake pour stocker les données de la table Delta dans Amazon S3 après l'opération de changement de nom de colonne. Pour plus d'informations sur cette fonctionnalité expérimentale dans Delta Lake, consultez Opération de changement de nom de colonne dans le guide de l'utilisateur Delta Lake.
- La version 6.11.0 résout un problème qui peut survenir lorsque vous créez un nœud périphérique en répliquant l'un des nœuds primaires à partir d'un cluster comportant plusieurs nœuds primaires. Le nœud périphérique répliqué peut retarder les opérations de réduction d'échelle ou entraîner une

utilisation élevée de la mémoire sur les nœuds primaires. Pour plus d'informations sur la création d'un nœud périphérique pour communiquer avec votre cluster EMR, consultez la section Edge Node Creator dans le `aws-samples` référentiel sur. GitHub

- La version 6.11.0 améliore le processus d'automatisation utilisé par Amazon EMR pour remonter les volumes Amazon EBS sur une instance après un redémarrage.
- La version 6.11.0 corrige un problème qui entraînait des écarts intermittents dans les métriques Hadoop publiées par Amazon EMR sur Amazon. CloudWatch
- La version 6.11.0 résout un problème lié aux clusters EMR où une mise à jour du fichier de configuration YARN contenant la liste d'exclusion des nœuds du cluster est interrompue en raison d'une surutilisation du disque. La mise à jour incomplète entrave les futures opérations de réduction de la taille du cluster. Cette version garantit que votre cluster reste sain et que les opérations de dimensionnement fonctionnent comme prévu.
- La taille du volume racine par défaut est passée à 15 Go dans Amazon EMR 6.10.0 et versions ultérieures. Les versions antérieures ont une taille de volume racine par défaut de 10 Go.
- Hadoop 3.3.3 a introduit une modification dans YARN (YARN-9608) qui maintient les nœuds sur lesquels les conteneurs s'exécutaient dans un état de mise hors service jusqu'à ce que l'application soit terminée. Cette modification permet de s'assurer que les données locales telles que les données réorganisées ne sont pas perdues et que vous n'avez pas besoin de réexécuter la tâche. Cette approche peut également entraîner une sous-utilisation des ressources sur les clusters avec ou sans activation de la mise à l'échelle gérée.

  Dans les versions 6.11.0 et supérieures d'Amazon EMR ainsi que dans les versions 6.8.1, 6.9.1 et 6.10.1, la valeur de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` est définie sur `false` dans `yarn-site.xml` pour résoudre ce problème.

  Bien que ce correctif règle les problèmes introduits par YARN-9608, il peut entraîner l'échec des tâches Hive en raison de la perte de données de réorganisation sur les clusters pour lesquels la mise à l'échelle gérée est activée. Nous avons atténué ce risque dans cette version en définissant également `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` pour les charges de travail Hive. Cette configuration n'est disponible qu'à partir de la version 6.11.0 d'Amazon EMR.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

> **Note**
>
> Cette version ne bénéficie plus de mises à jour automatiques de l'AMI puisqu'elle a été suivie d'une version supplémentaire de correctifs. La version du correctif est indiquée par le numéro qui suit la deuxième décimale (`6.8.1`). Pour savoir si vous utilisez la dernière version du correctif, consultez les versions disponibles dans le Guide des versions, ou consultez le menu déroulant des versions d'Amazon EMR lorsque vous créez un cluster dans la console, ou utilisez l'action d'API `ListReleaseLabels` ou de CLI `list-release-labels`. Pour être tenu au courant des nouvelles versions, abonnez-vous au flux RSS sur la page Quoi de neuf ?

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Zurich), Europe (Milan), Europe (Espagne), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Zurich), Europe (Milan), Europe (Espagne), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre)

Version 6.10.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.10.0. Les modifications ont été apportées à la version 6.9.0. Pour plus d'informations sur le calendrier de publication, consultez le journal des modifications.

Nouvelles fonctionnalités

- Amazon EMR 6.10.0 prend en charge Apache Spark 3.3.1, Apache Spark RAPIDS 22.12.0, CUDA 11.8.0, Apache Hudi 0.12.2-amzn-0, Apache Iceberg 1.1.0-amzn-0, Trino 403 et PrestoDB 0.278.1.
- Amazon EMR 6.10.0 inclut un connecteur Trino-Hudi natif qui fournit un accès en lecture aux données des tables Hudi. Vous pouvez activer le connecteur avec `trino-cli --catalog hudi` et le configurer en fonction de vos besoins avec `trino-connector-hudi`. L'intégration native avec Amazon EMR signifie que vous n'avez plus besoin d'utiliser `trino-connector-hive` pour interroger les tables Hudi. Pour obtenir la liste des configurations prises en charge avec le nouveau connecteur, consultez la page Connecteur Hudi dans la documentation de Trino.
- Les versions 6.10.0 et supérieures d'Amazon EMR prennent en charge l'intégration d'Apache Zeppelin avec Apache Flink. Pour plus d’informations, consultez Travailler avec les jobs Flink de Zeppelin dans Amazon EMR.

Problèmes connus

- Hadoop 3.3.3 a introduit une modification dans YARN (YARN-9608) qui maintient les nœuds sur lesquels les conteneurs s'exécutaient dans un état de mise hors service jusqu'à ce que l'application soit terminée. Cette modification permet de s'assurer que les données locales telles que les données réorganisées ne sont pas perdues et que vous n'avez pas besoin de réexécuter la tâche. Cette approche peut également entraîner une sous-utilisation des ressources sur les clusters avec ou sans activation de la mise à l'échelle gérée.

  Pour contourner ce problème dans Amazon EMR 6.10.0, vous pouvez définir la valeur de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` sur `false` dans `yarn-site.xml`. Dans les versions 6.11.0 et supérieures d'Amazon EMR, ainsi que 6.8.1, 6.9.1 et 6.10.1, la configuration est définie sur `false` par défaut pour résoudre ce problème.

Modifications, améliorations et problèmes résolus

- Amazon EMR 6.10.0 supprime la dépendance sur `minimal-json.jar` pour l'intégration Amazon Redshift pour Apache Spark et ajoute automatiquement les fichiers JAR liés à Spark-Redshift requis au chemin de classe de l'exécuteur pour Spark : `spark-redshift.jar`, `spark-avro.jar` et `RedshiftJDBC.jar`.
- La version 6.10.0 améliore le démon de gestion des journaux sur le cluster afin de surveiller des dossiers de journaux supplémentaires dans votre cluster EMR. Cette amélioration permet de minimiser les scénarios de surutilisation des disques.
- La version 6.10.0 redémarre automatiquement le démon de gestion des journaux du cluster lorsqu'il s'arrête. Cette amélioration réduit le risque que les nœuds apparaissent défectueux en raison d'une surutilisation du disque.
- Amazon EMR 6.10.0 prend en charge les points de terminaison régionaux pour le mappage des utilisateurs EMRFS.
- La taille du volume racine par défaut est passée à 15 Go dans Amazon EMR 6.10.0 et versions ultérieures. Les versions antérieures ont une taille de volume racine par défaut de 10 Go.
- La version 6.10.0 corrige un problème qui provoquait le blocage des tâches Spark lorsque tous les exécuteurs Spark restants se trouvaient sur un hôte en cours de mise hors service avec le gestionnaire de ressources YARN.
- Avec Amazon EMR 6.6.0 à 6.9.x, les requêtes INSERT avec partition dynamique et clause ORDER BY ou SORT BY auront toujours deux réducteurs. Ce problème est dû à la modification d'OSS HIVE-20703, qui place l'optimisation des partitions dynamiques de tri dans le cadre d'une décision basée sur les coûts. Si votre charge de travail ne nécessite pas le tri des partitions dynamiques, nous vous recommandons de définir la propriété `hive.optimize.sort.dynamic.partition.threshold` sur `-1` pour désactiver la nouvelle fonctionnalité et obtenir le nombre de réducteurs correctement calculé. Ce problème est résolu dans OSS Hive dans le cadre de HIVE-22269 et dans Amazon EMR 6.10.0.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

> **Note**
>
> Cette version ne bénéficie plus de mises à jour automatiques de l'AMI puisqu'elle a été suivie d'une version supplémentaire de correctifs. La version du correctif est indiquée par

le numéro qui suit la deuxième décimale (`6.8.1`). Pour savoir si vous utilisez la dernière version du correctif, consultez les versions disponibles dans le Guide des versions, ou consultez le menu déroulant des versions d'Amazon EMR lorsque vous créez un cluster dans la console, ou utilisez l'action d'API `ListReleaseLabels` ou de CLI `list-release-labels`. Pour être tenu au courant des nouvelles versions, abonnez-vous au flux RSS sur la page Quoi de neuf ?

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada Ouest (Calgary), Europe (Zurich)
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Zurich), Europe (Milan), Europe (Espagne), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Zurich), Europe (Milan), Europe (Espagne), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Zurich), Europe (Milan), Europe (Espagne), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Hyderaba d), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 418,0	4,14.311	3 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Zurich), Europe (Milan), Europe (Espagne), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 404.1	4,14.311	18 avril 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)
2,0.2023 404.0	4,14.311	10 avril 2023	USA Est (Virginie du Nord), Europe (Paris)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 320,0	4,14,309	30 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 207.0	4,14,304	22 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU)

Version 6.9.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.9.0. Il s'agit des modifications apportées à la version Amazon EMR 6.8.0. Pour plus d'informations sur le calendrier de publication, consultez le journal des modifications.

Nouvelles fonctions

- La version 6.9.0 d'Amazon EMR prend en charge Apache Spark RAPIDS 22.08.0, Apache Hudi 0.12.1, Apache Iceberg 0.14.1, Trino 398 et Tez 0.10.2.
- La version 6.9.0 d'Amazon EMR inclut une nouvelle application open source, Delta Lake 2.1.0.
- L'intégration d'Amazon Redshift à Apache Spark est incluse dans les versions 6.9.0 et ultérieures d'Amazon EMR. Auparavant un outil open-source, l'intégration native est un connecteur Spark que vous pouvez utiliser pour créer des applications Apache Spark capables de lire et d'écrire des données sur Amazon Redshift et Amazon Redshift sans serveur. Pour de plus amples informations, veuillez consulter Utilisation de l'intégration d'Amazon Redshift pour Apache Spark avec Amazon EMR .
- La version 6.9.0 d'Amazon EMR ajoute la prise en charge de l'archivage des journaux dans Amazon S3 lors de la réduction de la taille du cluster. Auparavant, vous pouviez uniquement archiver les fichiers journaux sur Amazon S3 lors de la résiliation du cluster. Cette nouvelle fonctionnalité garantit que les fichiers journaux générés sur le cluster sont conservés sur Amazon S3 même après la résiliation du nœud. Pour plus d'informations, consultez Configuration de la journalisation et du débogage de cluster.
- Pour prendre en charge les requêtes de longue durée, Trino inclut désormais un mécanisme d'exécution tolérant aux pannes. L'exécution tolérante aux pannes atténue les échecs des requêtes en réessayant les requêtes qui ont échoué ou les tâches correspondantes.
- Vous pouvez utiliser Apache Flink sur Amazon EMR pour le traitement `BATCH` et `STREAM` unifié des tables Apache Hive ou des métadonnées de n'importe quelle source de tables Flink telle que Iceberg, Kinesis ou Kafka. Vous pouvez spécifier le catalogue de données AWS Glue comme métastore pour Flink à l'aide de l'API AWS Management Console, AWS CLI ou Amazon EMR. Pour de plus amples informations, veuillez consulter Configuration de Flink dans Amazon EMR.
- Vous pouvez désormais spécifier des rôles d'exécution AWS Identity and Access Management (IAM) et un contrôle d'accès AWS Lake Formation basé pour les requêtes Apache Spark, Apache Hive et Presto sur Amazon EMR sur des clusters EC2 avec Amazon AI Studio. SageMaker Pour obtenir des informations supplémentaires, consultez Configuration des rôles d'exécution pour les étapes d'Amazon EMR.

Problèmes connus

- Pour Amazon EMR version 6.9.0, Trino ne fonctionne pas sur les clusters activés pour Apache Ranger. Si vous devez utiliser Trino avec Ranger, contactez Support.

- Si vous utilisez l'intégration Amazon Redshift à Apache Spark et que vous disposez d'un champ de type heure, timetz, horodatage ou timestamptz avec une précision de l'ordre de la microseconde au format Parquet, le connecteur arrondit les valeurs temporelles à la milliseconde la plus proche. Pour contourner le problème, utilisez le paramètre `unload_s3_format` de format de déchargement du texte.
- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.
- Les connexions aux clusters Amazon EMR depuis Amazon SageMaker AI Studio peuvent échouer par intermittence avec un code de réponse 403 Forbidden. Cette erreur se produit lorsque la configuration du rôle IAM sur le cluster prend plus de 60 secondes. Pour contourner le problème, vous pouvez installer un correctif Amazon EMR pour activer les nouvelles tentatives et augmenter le délai d'expiration à un minimum de 300 secondes. Suivez les étapes ci-dessous pour appliquer l'action d'amorçage lorsque vous lancez votre cluster.

  1. Téléchargez le script bootstrap et les fichiers RPM depuis l'Amazon S3 URIs suivant.

     ```
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/replace-rpms.sh
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/emr-secret-agent-1.18.0-SNAPSHOT20221121212949.noarch.rpm
     ```

2. Chargez les fichiers de l'étape précédente sur un compartiment Amazon S3 dont vous êtes propriétaire. Le compartiment doit se trouver à l' Région AWS endroit où vous prévoyez de lancer le cluster.
3. Incluez l'action d'amorçage suivante lorsque vous lancez votre cluster EMR. Remplacez *bootstrap_URI* et *RPM_URI* par le correspondant URIs provenant d'Amazon S3.

```
--bootstrap-actions "Path=bootstrap_URI,Args=[RPM_URI]"
```

- Avec les versions 5.36.0 et 6.6.0 à 6.9.0 d'Amazon EMR, les composants de service `SecretAgent` et `RecordServer` peuvent subir une perte de données de journal en raison d'une configuration incorrecte du modèle de nom de fichier dans les propriétés de Log4j2. En cas de configuration incorrecte, les composants ne génèrent qu'un seul fichier journal par jour. Lorsque la stratégie de rotation est appliquée, elle remplace le fichier existant au lieu de générer un nouveau fichier journal comme prévu. Pour contourner le problème, utilisez une action d'amorçage pour générer des journaux toutes les heures et ajoutez un nombre entier auto-incrémenté dans le nom du fichier pour gérer la rotation.

  Pour les versions 6.6.0 à 6.9.0 d'Amazon EMR, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
  ```

  Pour Amazon EMR 5.36.0, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
  ```

- Apache Flink fournit des FileSystem connecteurs natifs S3 FileSystem et Hadoop, qui permettent aux applications de créer FileSink et d'écrire les données dans Amazon S3. Cela FileSink échoue avec l'une des deux exceptions suivantes.

  ```
  java.lang.UnsupportedOperationException: Recoverable writers on Hadoop are only
   supported for HDFS
  ```

  ```
  Caused by: java.lang.NoSuchMethodError:
   org.apache.hadoop.io.retry.RetryPolicies.retryOtherThanRemoteAndSaslException(Lorg/
  ```

```
apache/hadoop/io/retry/RetryPolicy;Ljava/util/Map;)Lorg/apache/hadoop/io/retry/
RetryPolicy;
                                        at
 org.apache.hadoop.yarn.client.RMProxy.createRetryPolicy(RMProxy.java:302) ~[hadoop-
yarn-common-3.3.3-amzn-0.jar:?]
```

Pour contourner le problème, vous pouvez installer un correctif Amazon EMR, qui corrige le problème ci-dessus dans Flink. Suivez les étapes suivantes pour appliquer l'action d'amorçage lors du lancement de votre cluster.

1. Téléchargez le fichier flink-rpm dans votre compartiment Amazon S3. Votre chemin RPM est `s3://DOC-EXAMPLE-BUCKET/rpms/flink/`.
2. Téléchargez le script d'amorçage et les fichiers RPM depuis Amazon S3 en utilisant l'URI suivant. Remplacez *regionName* par l' Région AWS endroit où vous prévoyez de lancer le cluster.

   ```
   s3://emr-data-access-control-regionName/customer-bootstrap-actions/gcsc/replace-
   rpms.sh
   ```

3. Hadoop 3.3.3 a introduit une modification dans YARN (YARN-9608) qui maintient les nœuds sur lesquels les conteneurs s'exécutaient dans un état de mise hors service jusqu'à ce que l'application soit terminée. Cette modification permet de s'assurer que les données locales telles que les données réorganisées ne sont pas perdues et que vous n'avez pas besoin de réexécuter la tâche. Dans Amazon EMR 6.8.0 et 6.9.0, cette approche peut également entraîner une sous-utilisation des ressources sur les clusters avec ou sans activation de la mise à l'échelle gérée.

   Avec Amazon EMR 6.10.0, il existe une solution à ce problème qui consiste à définir la valeur de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` sur `false` dans `yarn-site.xml`. Dans les versions 6.11.0 et supérieures d'Amazon EMR, ainsi que 6.8.1, 6.9.1 et 6.10.1, la configuration est définie sur `false` par défaut pour résoudre ce problème.

Modifications, améliorations et problèmes résolus

- Pour Amazon EMR version 6.9.0 et versions ultérieures, tous les composants installés par Amazon EMR qui utilisent les bibliothèques Log4j utilisent Log4j version 2.17.1 ou ultérieure.
- Lorsque vous utilisez le connecteur DynamoDB avec Spark sur les versions 6.6.0, 6.7.0 et 6.8.0 d'Amazon EMR, toutes les lectures de votre table renvoient un résultat vide, même si la division

d'entrée fait référence à des données non vides. La version 6.9.0 d'Amazon EMR résout ce problème.

- Amazon EMR 6.9.0 ajoute une prise en charge limitée du contrôle d'accès basé sur Lake Formation avec Apache Hudi lors de la lecture de données à l'aide de Spark SQL. La prise en charge concerne les requêtes SELECT utilisant Spark SQL et se limite au contrôle d'accès au niveau des colonnes. Pour plus d'informations, consultez Hudi et Lake Formation.
- Lorsque vous utilisez Amazon EMR 6.9.0 pour créer un cluster Hadoop avec les étiquettes de nœuds activées, l'API de métriques YARN renvoie des informations agrégées sur toutes les partitions, au lieu de la partition par défaut. Pour plus d'informations, consultez YARN-11414.
- Avec la version 6.9.0 d'Amazon EMR, nous avons mis à jour Trino vers la version 398, qui utilise Java 17. La version précédemment prise en charge de Trino pour Amazon EMR 6.8.0 était Trino 388 fonctionnant sous Java 11. Pour plus d'informations sur cette modification, consultez Mises à jour de Trino vers Java 17 sur le blog de Trino.
- Cette version corrige un problème d'incompatibilité de séquence temporelle entre Apache BigTop et Amazon EMR lors de la séquence de démarrage EC2 du cluster. Ce décalage se produit lorsqu'un système tente d'effectuer deux ou plusieurs opérations en même temps au lieu de les effectuer dans le bon ordre. Par conséquent, certaines configurations de cluster ont connu des délais de démarrage des instances et des temps de démarrage des clusters plus lents.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

> **Note**
>
> Cette version ne bénéficie plus de mises à jour automatiques de l'AMI puisqu'elle a été suivie d'une version supplémentaire de correctifs. La version du correctif est indiquée par le numéro qui suit la deuxième décimale (`6.8.1`). Pour savoir si vous utilisez la dernière version du correctif, consultez les versions disponibles dans le Guide des versions, ou consultez le menu déroulant des versions d'Amazon EMR lorsque vous créez un cluster dans la console, ou utilisez l'action d'API `ListReleaseLabels` ou de CLI `list-release-labels`. Pour être tenu au courant des nouvelles versions, abonnez-vous au flux RSS sur la page Quoi de neuf ?

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Aviv), Canada (centre), Canada Ouest (Calgary), Europe (Espagne), Chine (Ningxia), Europe (Zurich)
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), Canada Ouest (Calgary), AWS GovCloud (USA Est),

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			AWS GovCloud (USA Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Europe (Zurich)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Canada Ouest (Calgary)
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre), Israël (Tel Aviv)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 418,0	4,14.311	3 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 404.1	4,14.311	18 avril 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0.2023 404.0	4,14.311	10 avril 2023	USA Est (Virginie du Nord), Europe (Paris)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 320,0	4,14,309	30 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 307.0	4,14,305	15 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 207.0	4,14,304	22 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2022 210.1	4,14,301	12 janvier 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 103.3	4,14.296	5 décembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

Version 6.8.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.8.0. Les modifications ont été apportées à la version 6.7.0.

Nouvelles fonctions

- La fonctionnalité Amazon EMR Steps prend désormais en charge le point de terminaison et les clients Apache Livy. JDBC/ODBC Pour obtenir des informations supplémentaires, consultez Configuration des rôles d'exécution pour les étapes d'Amazon EMR.
- La version 6.8.0 d'Amazon EMR est fournie avec la version 2.4.12 d'Apache HBase . Avec cette HBase version, vous pouvez à la fois archiver et supprimer vos HBase tables. Le processus d'archivage Amazon S3 renomme tous les fichiers de table dans le répertoire d'archive. Ce processus peut être long et coûteux. Vous pouvez désormais ignorer le processus d'archivage et supprimer rapidement des tables volumineuses. Pour de plus amples informations, veuillez consulter Utilisation de la HBase coque.

Problèmes connus

- Hadoop 3.3.3 a introduit une modification dans YARN (YARN-9608) qui maintient les nœuds sur lesquels les conteneurs s'exécutaient dans un état de mise hors service jusqu'à ce que l'application soit terminée. Cette modification permet de s'assurer que les données locales telles que les données réorganisées ne sont pas perdues et que vous n'avez pas besoin de réexécuter la tâche. Dans Amazon EMR 6.8.0 et 6.9.0, cette approche peut également entraîner une sous-utilisation des ressources sur les clusters avec ou sans activation de la mise à l'échelle gérée.

  Avec Amazon EMR 6.10.0, il existe une solution à ce problème qui consiste à définir la valeur de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` sur `false` dans `yarn-site.xml`. Dans les versions 6.11.0 et supérieures d'Amazon EMR, ainsi que 6.8.1, 6.9.1 et 6.10.1, la configuration est définie sur `false` par défaut pour résoudre ce problème.

Modifications, améliorations et problèmes résolus

- Lorsqu'Amazon EMR version 6.5.0, 6.6.0 ou 6.7.0 lisait les tables Apache Phoenix via le shell Apache Spark, Amazon EMR a produit une erreur `NoSuchMethodError`. La version 6.8.0 d'Amazon EMR résout ce problème.
- La version 6.8.0 d'Amazon EMR est fournie avec Apache Hudi 0.11.1 ; toutefois, les clusters Amazon EMR 6.8.0 sont également compatibles avec le `hudi-spark3.3-bundle_2.12` open source de Hudi 0.12.0.

- La version 6.8.0 d'Amazon EMR est fournie avec la version 3.3.0 d'Apache Spark. Cette version de Spark utilise Apache Log4j 2 et le fichier `log4j2.properties` pour configurer Log4j dans les processus Spark. Si vous utilisez Spark dans le cluster ou si vous créez des clusters EMR avec des paramètres de configuration personnalisés, et que vous voulez passer à la version 6.8.0 d'Amazon EMR, vous devez migrer vers la nouvelle classification de configuration `spark-log4j2` et le nouveau format de clé pour Apache Log4j 2. Pour de plus amples informations, veuillez consulter Migration d'Apache Log4j 1.x vers Log4j 2.x.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

> **Note**
>
> Cette version ne bénéficie plus de mises à jour automatiques de l'AMI puisqu'elle a été suivie d'une version supplémentaire de correctifs. La version du correctif est indiquée par le numéro qui suit la deuxième décimale (`6.8.1`). Pour savoir si vous utilisez la dernière version du correctif, consultez les versions disponibles dans le Guide des versions, ou consultez le menu déroulant des versions d'Amazon EMR lorsque vous créez un cluster dans la console, ou utilisez l'action d'API `ListReleaseLabels` ou de CLI `list-release-labels`. Pour être tenu au courant des nouvelles versions, abonnez-vous au flux RSS sur la page Quoi de neuf ?

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Pacifique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Ohio), Afrique (Le Cap), Europe (Irlande), Moyen-Orient (Émirats arabes unis), Europe (Francfor t), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), USA Est (N. Virginie), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Asie-Pacifique (Melbourne), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (AWS GovCloud Singapour), Asie-Paci fique (Sydney), Israël (Tel Aviv), Canada (centre), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Moyen-Ori ent (Émirats arabes unis), Canada (Centre), AWS GovCloud (USA Est), AWS GovCloud (USA

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Ouest), Chine (Pékin), Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Asie-Pacifique (Melbourne), Israël (Tel Aviv), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe (Espagne), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Zurich), Asie-Pacifique (Melbourne), Israël (Tel Aviv)
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Asie-Pacifique (Hong Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta), Asie-Pacifique (Melbourn e), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 418,0	4,14.311	3 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 404.1	4,14.311	18 avril 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0.2023 404.0	4,14.311	10 avril 2023	USA Est (Virginie du Nord), Europe (Paris)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 320,0	4,14,309	30 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 307.0	4,14,305	15 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 207.0	4,14,304	22 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 119.1	4,14,301	3 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2022 210.1	4,14,301	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 103.3	4,14.296	5 décembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 004.0	4,14.294	2 novembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 912.1	4,14,291	6 septembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

Problèmes connus

- Lorsque vous utilisez le connecteur DynamoDB avec Spark sur les versions 6.6.0, 6.7.0 et 6.8.0 d'Amazon EMR, toutes les lectures de votre table renvoient un résultat vide, même si la division d'entrée fait référence à des données non vides. Cela est dû au fait que Spark 3.2.0 définit `spark.hadoopRDD.ignoreEmptySplits` sur `true` par défaut. Pour contourner le problème,

définissez explicitement `spark.hadoopRDD.ignoreEmptySplits` sur `false`. La version 6.9.0 d'Amazon EMR résout ce problème.

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.
- Avec les versions 5.36.0 et 6.6.0 à 6.9.0 d'Amazon EMR, les composants de service `SecretAgent` et `RecordServer` peuvent subir une perte de données de journal en raison d'une configuration incorrecte du modèle de nom de fichier dans les propriétés de Log4j2. En cas de configuration incorrecte, les composants ne génèrent qu'un seul fichier journal par jour. Lorsque la stratégie de rotation est appliquée, elle remplace le fichier existant au lieu de générer un nouveau fichier journal comme prévu. Pour contourner le problème, utilisez une action d'amorçage pour générer des journaux toutes les heures et ajoutez un nombre entier auto-incrémenté dans le nom du fichier pour gérer la rotation.

  Pour les versions 6.6.0 à 6.9.0 d'Amazon EMR, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
  ```

  Pour Amazon EMR 5.36.0, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Pour plus d'informations sur le calendrier de publication, consultez le journal des modifications.

Version 6.7.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.7.0. Les modifications ont été apportées à la version 6.6.0.

Date de parution initiale : 15 juillet 2022

Nouvelles fonctions

- Amazon EMR prend désormais en charge Apache Spark 3.2.1, Apache Hive 3.1.3, HUDI 0.11, PrestoDB 0.272 et Trino 0.378.
- Prend en charge les contrôles d'accès basés sur les rôles IAM et les formations lacustres avec des étapes EMR (Spark, Hive) pour Amazon EMR sur des clusters. EC2
- Prend en charge les instructions de définition de données Apache Spark sur les clusters compatibles avec Apache Ranger. Cela inclut désormais la prise en charge des applications Trino lisant et écrivant des métadonnées Apache Hive sur des clusters compatibles avec Apache Ranger. Pour plus d'informations, consultez Mise en place d'une gouvernance fédérée à l'aide de Trino et d'Apache Ranger sur Amazon EMR.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn),

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est (Ohio), Afrique (Le Cap), Europe (Irlande) , Moyen-Orient (Émirats arabes unis), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), Asie-Pacifique (Hyderaba d), USA Est (N. Virginie) , Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres), Europe (Milan), Asie-Pacifique (Tokyo), (États-Unis est), (États-Unis ouest) AWS GovCloud , ouest des États-Unis (Oregon), ouest des États-Unis (Californ ie du Nord), Asie-Paci fique (Singapour), Asie-Pacifique (Sydney) AWS GovCloud), Israël (Tel Aviv), Canada (centre), Europe (Espagne), Chine (Ningxia), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Espagne), Europe (Stockholm), Europe (Zurich), Asie-Pacifique (Hong Kong), Asie-Paci fique (Hyderabad), Asie-Pacifique (Jakarta), Asie-Pacifique (Mumbai), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (Émirats arabes unis), Canada (Centre), (États-Unis- AWS GovCloud Est), AWS GovCloud (US-Ouest), Chine (Pékin),

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			Chine (Ningxia), Israël (Tel Aviv)
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Moyen-Orient Est (Émirats arabes unis), Israël (Tel Aviv), Europe (Zurich)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia), Asie Pacifique (Hyderabad), Moyen-Orient (Émirats arabes unis), Europe

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Espagne), Europe (Zurich)
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 822.0	4,1,4322	30 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 418,0	4,14.311	3 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 404.1	4,14.311	18 avril 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0.2023 404.0	4,14.311	10 avril 2023	USA Est (Virginie du Nord), Europe (Paris)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 320,0	4,14,309	30 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 307.0	4,14,305	15 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 207.0	4,14,304	22 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 119.1	4,14,301	3 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2022 210.1	4,14,301	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 103.3	4,14.296	5 décembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 004.0	4,14.294	2 novembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 912.1	4,14,291	7 octobre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0,2022 719,0	4,14.287	10 août 2022	`us-west-1` , `eu-west-3` , `eu-north-1` , `ap-south-1` , `me-south-1`

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 606.1	4,14,281	15 juillet 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

Problèmes connus

- Lorsque les versions 6.5.0, 6.6.0 ou 6.7.0 d'Amazon EMR lisent les tables Apache Phoenix via le shell Apache Spark, une erreur `NoSuchMethodError` se produit car Amazon EMR utilise une `Hbase.compat.version` incorrecte. La version 6.8.0 d'Amazon EMR résout ce problème.
- Lorsque vous utilisez le connecteur DynamoDB avec Spark sur les versions 6.6.0, 6.7.0 et 6.8.0 d'Amazon EMR, toutes les lectures de votre table renvoient un résultat vide, même si la division

d'entrée fait référence à des données non vides. Cela est dû au fait que Spark 3.2.0 définit `spark.hadoopRDD.ignoreEmptySplits` sur `true` par défaut. Pour contourner le problème, définissez explicitement `spark.hadoopRDD.ignoreEmptySplits` sur `false`. La version 6.9.0 d'Amazon EMR résout ce problème.

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.
- Avec les versions 5.36.0 et 6.6.0 à 6.9.0 d'Amazon EMR, les composants de service `SecretAgent` et `RecordServer` peuvent subir une perte de données de journal en raison d'une configuration incorrecte du modèle de nom de fichier dans les propriétés de Log4j2. En cas de configuration incorrecte, les composants ne génèrent qu'un seul fichier journal par jour. Lorsque la stratégie de rotation est appliquée, elle remplace le fichier existant au lieu de générer un nouveau fichier journal comme prévu. Pour contourner le problème, utilisez une action d'amorçage pour générer des journaux toutes les heures et ajoutez un nombre entier auto-incrémenté dans le nom du fichier pour gérer la rotation.

  Pour les versions 6.6.0 à 6.9.0 d'Amazon EMR, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Pour Amazon EMR 5.36.0, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- L'API `GetClusterSessionCredentials` n'est pas prise en charge avec les clusters qui s'exécutent sur Amazon EMR 6.7 ou version antérieure.

Version 6.6.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.6.0. Les modifications ont été apportées à la version 6.5.0.

Date de parution initiale : 9 mai 2022

Date de mise à jour de la documentation : 15 juin 2022

Nouvelles fonctions

- Amazon EMR 6.6 prend désormais en charge Apache Spark 3.2, Apache Spark RAPIDS 22.02, CUDA 11, Apache Hudi 0.10.1, Apache Iceberg 0.13, Trino 0.367 et PrestoDB 0.267.
- Lorsque vous lancez un cluster avec le dernier correctif d'Amazon EMR 5.36 ou supérieur, 6.6 ou supérieur, ou 7.0 ou supérieur, Amazon EMR utilise la dernière version d'Amazon Linux 2023 ou Amazon Linux 2 pour l'AMI Amazon EMR par défaut. Pour plus d'informations, consultez Utilisation de l'AMI Amazon Linux par défaut pour Amazon EMR.

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 527.1	4.14.355-277.647 .amzn2	19 juin 2025	Europe (Stockholm), Moyen-Orient (Bahreïn), Chine (Pékin), Asie-Paci fique (Mumbai), Europe (Paris), Asie-Pacifique (Jakarta), États-Unis Est

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
			(Ohio), Afrique (Le Cap), Europe (Irlande), Europe (Francfort), Amérique du Sud (São Paulo), Asie-Pacifique (Hong Kong), USA Est (Virginie du Nord), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Europe (Londres) , Europe (Milan), Asie-Paci fique (Tokyo), (États-Un is est) AWS GovCloud , (États-Unis ouest), ouest des États-Unis (Oregon), ouest des États-Unis (Californie du Nord), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Canada (centre), Chine (Ningxia) AWS GovCloud

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 512,0	4.14.355-277.643.amzn2	4 juin 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 428,0	4,14,355-276,639. amzn2	23 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Francfort), Europe (Irlande), Europe (Londres), Europe (Milan), Europe (Paris), Europe (Stockholm), Asie-Paci fique (Hong Kong), Asie-Pacifique (Jakarta), Asie-Pacifique (Osaka), Asie-Pacifique (Séoul), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Tokyo), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Ori ent (Bahreïn), Canada (centre), (États-Unis est), (ouest des États-Unis AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 414,0	4,14,355-276,618. amzn2	12 mai 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 321.0	4,14,355	9 avril 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 305.0	4,14,355	18 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 220,0	4,14,355	8 mars 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 201.0	4,14,355	28 février 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2025 123.4	4,14,355	27 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2025 116,0	4,14,355	23 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 217,0	4,14,355	8 janvier 2025	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 001.0	4,14,352	4 octobre 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 816,0	4,14,350	21 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 809.0	4,1,4349	20 août 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2024 719,0	4,14.348	25 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 709.1	4,14.348	23 juillet 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 223.0	4,14.336	8 mars 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 131.0	4,14.336	14 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 124.0	4,14.336	7 février 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2024 109.0	4,1,4334	24 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 218,0	4,14,330	2 janvier 2024	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 206,0	4,14,330	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 116,0	4,14,328	11 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 101,0	4,1,4327	16 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 020.1	4,1,4326	7 novembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 012.1	4,1,4326	26 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 926,0	4,1,4322	19 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong Kong), Asie-Paci fique (Tokyo), Asie-Paci fique (Séoul), Asie-Paci fique (Osaka), Asie-Pacifique (Singapour) Singapour), Asie-Paci fique (Sydney), Asie-Paci fique (Jakarta), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (centre), (États-Unis ouest), (États-Unis est AWS GovCloud), Chine (Pékin AWS GovCloud), Chine (Ningxia)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 906.0	4,1,4322	4 octobre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 822.0	4,1,4322	30 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 808.0	4,14,320	24 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 727,0	4,14,320	14 août 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Francfor t), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Paci fique (Mumbai), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 719,0	4,14,320	02/08/2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Europe (Stockholm), Europe (Milan), Europe (Espagne) , Europe (Francfor t), Europe (Zurich), Europe (Irlande), Europe (Londres), Europe (Paris), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Paci fique (Hyderabad), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Asie-Pacifique (Jakarta) , Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn), Moyen-Orient (EAU), Canada (Centre)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 628,0	4,14,318	12 juillet 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 612,0	4,1,4314	23 juin 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 504.1	4,14.313	16 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 418,0	4,14.311	3 mai 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 404.1	4,14.311	18 avril 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0.2023 404.0	4,14.311	10 avril 2023	USA Est (Virginie du Nord), Europe (Paris)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2023 320,0	4,14,309	30 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 307.0	4,14,305	15 mars 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2023 207.0	4,14,304	22 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2023 119.1	4,14,301	3 février 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2.0.2022 210.1	4,14,301	22 décembre 2023	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockhol m), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfor t), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Paci fique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 103.3	4,14.296	5 décembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 004.0	4,14.294	2 novembre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 912.1	4,14,291	7 octobre 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)
2,0,2022 805.0	4,14.287	30 août 2022	`us-west-1`

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0,2022 719,0	4,14.287	10 août 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 426.0	4,14,281	10 juin 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

OsRelea Label (Version Amazon Linux)	Version du noyau Amazon Linux	Date de disponibilité	Régions prises en charge
2,0.2022 406.1	4,14,275	2 mai 2022	USA Est (Virginie du Nord), USA Est (Ohio), USA Ouest (Californie du Nord), USA Ouest (Oregon), Canada (Centre), Europe (Stockholm), Europe (Irlande), Europe (Londres), Europe (Paris), Europe (Francfort), Europe (Milan), Asie-Pacifique (Hong-Kong), Asie-Pacifique (Mumbai), Asie-Pacifique (Jakarta), Asie-Pacifique (Tokyo), Asie-Pacifique (Séoul), Asie-Pacifique (Osaka), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Afrique (Le Cap), Amérique du Sud (São Paulo), Moyen-Orient (Bahreïn)

- Avec Amazon EMR 6.6 et les versions ultérieures, les applications qui utilisent Log4j 1.x et Log4j 2.x sont mises à niveau pour utiliser respectivement Log4j 1.2.17 (ou supérieur) et Log4j 2.17.1 (ou supérieur), et n'ont pas besoin d'utiliser les actions d'amorçage fournies pour atténuer les problèmes liés aux CVE.
- [Mise à l'échelle gérée] Optimisation de la mise à l'échelle gérée des données de réorganisation Spark – Pour Amazon EMR versions 5.34.0 et ultérieures, et EMR versions 6.4.0 et ultérieures, la mise à l'échelle gérée prend désormais en compte les données de réorganisation Spark (données que Spark redistribue entre les partitions pour effectuer des opérations spécifiques).

Pour plus d'informations sur les opérations de réorganisation, consultez Utilisation de la mise à l'échelle gérée par EMR dans Amazon EMR dans le Guide de gestion Amazon EMR et le Guide de programmation Spark.

- À partir d'Amazon EMR 5.32.0 et 6.5.0, le dimensionnement dynamique de l'exécuteur pour Apache Spark est activé par défaut. Pour activer ou désactiver cette fonctionnalité, vous pouvez utiliser le paramètre de configuration `spark.yarn.heterogeneousExecutors.enabled`.

Modifications, améliorations et problèmes résolus

- Amazon EMR réduit le temps de démarrage des clusters de 80 secondes en moyenne pour les clusters qui utilisent l'option AMI par défaut d'EMR et n'installent que des applications courantes, telles qu'Apache Hadoop, Apache Spark et Apache Hive.

Problèmes connus

- Lorsque les versions 6.5.0, 6.6.0 ou 6.7.0 d'Amazon EMR lisent les tables Apache Phoenix via le shell Apache Spark, une erreur `NoSuchMethodError` se produit car Amazon EMR utilise une `Hbase.compat.version` incorrecte. La version 6.8.0 d'Amazon EMR résout ce problème.
- Lorsque vous utilisez le connecteur DynamoDB avec Spark sur les versions 6.6.0, 6.7.0 et 6.8.0 d'Amazon EMR, toutes les lectures de votre table renvoient un résultat vide, même si la division d'entrée fait référence à des données non vides. Cela est dû au fait que Spark 3.2.0 définit `spark.hadoopRDD.ignoreEmptySplits` sur `true` par défaut. Pour contourner le problème, définissez explicitement `spark.hadoopRDD.ignoreEmptySplits` sur `false`. La version 6.9.0 d'Amazon EMR résout ce problème.
- Sur les clusters de longue durée de Trino, Amazon EMR 6.6.0 active les paramètres de journalisation du récupérateur de mémoire dans le fichier jvm.config de Trino afin d'obtenir de meilleures informations à partir des journaux du récupérateur de mémoire. Cette modification ajoute de nombreux journaux de collecte des déchets au fichier launcher.log (/var/log/trino/launcher.log). Si vous utilisez des clusters Trino dans Amazon EMR 6.6.0, vous pouvez rencontrer des nœuds à court d'espace disque après quelques jours d'exécution du cluster en raison des journaux ajoutés.

  La solution à ce problème consiste à exécuter le script ci-dessous en tant qu'action d'amorçage afin de désactiver les paramètres de journalisation du récupérateur de mémoire dans jvm.config lors de la création ou du clonage du cluster pour Amazon EMR 6.6.0.

```
#!/bin/bash
  set -ex
  PRESTO_PUPPET_DIR='/var/aws/emr/bigtop-deploy/puppet/modules/trino'
  sudo bash -c "sed -i '/-Xlog/d' ${PRESTO_PUPPET_DIR}/templates/jvm.config"
```

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.
- Avec les versions 5.36.0 et 6.6.0 à 6.9.0 d'Amazon EMR, les composants de service `SecretAgent` et `RecordServer` peuvent subir une perte de données de journal en raison d'une configuration incorrecte du modèle de nom de fichier dans les propriétés de Log4j2. En cas de configuration incorrecte, les composants ne génèrent qu'un seul fichier journal par jour. Lorsque la stratégie de rotation est appliquée, elle remplace le fichier existant au lieu de générer un nouveau fichier journal comme prévu. Pour contourner le problème, utilisez une action d'amorçage pour générer des journaux toutes les heures et ajoutez un nombre entier auto-incrémenté dans le nom du fichier pour gérer la rotation.

  Pour les versions 6.6.0 à 6.9.0 d'Amazon EMR, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Pour Amazon EMR 5.36.0, utilisez l'action de démarrage suivante lorsque vous lancez un cluster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Version 5.35.0

Ceci est la note de mise à jour de la version 5.35.0 d'Amazon EMR.

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.35.0. Les modifications ont été apportées à la version 5.34.0.

Date de parution initiale : 30 mars 2022

Nouvelles fonctions

- Les applications Amazon EMR version 5.35 qui utilisent Log4j 1.x et Log4j 2.x sont mises à niveau pour utiliser Log4j 1.2.17 (ou version ultérieure) et Log4j 2.17.1 (ou version ultérieure) respectivement, et ne nécessitent pas d'actions d'amorçage pour atténuer les problèmes CVE dans les versions précédentes. Consultez Approche visant à atténuer le CVE-2021-44228.

Modifications, améliorations et problèmes résolus

Modifications apportées à Flink

Type de modification	Description
Mises à niveau	• Mise à jour de la version de flink vers la version 1.14.2. • log4j a été mis à jour vers la version 2.17.1.

Modifications apportées à Hadoop

Type de modification	Description
Rétroportages de Hadoop open source depuis EMR 5.34.0	• YARN-10438 : Gérer un identifiant de conteneur nul dans le numéro du client () RMService getContainerReport • YARN-7266 : les threads du gestionnaire d'événements du serveur de chronologie sont verrouillés • YARN-10438 : ATS 1.5 ne démarre pas si les RollingLevelDb fichiers sont corrompus ou manquants • HADOOP-13500 : Synchronisation de l'itérati on de l'objet Propriétés de configuration • YARN-10651 : CapacityScheduler s'est écrasé avec NPE dedans. AbstractY arnScheduler updateNodeResource() • HDFS-12221 : Remplacer les caractères dans XmlEditsVisitor • HDFS-16410 : Analyse XML non sécurisée OfflineEditsXmlLoader
Modifications et corrections apportées à Hadoop	• Tomcat utilisé dans KMS et HttpFS est mis à niveau vers la version 8.5.75 • Dans la FileSystemOptimizedCommitter version 2, le marqueur de réussite était écrit dans le chemin de sortie CommitJob défini lors de la création du commit. Étant donné que les chemins de sortie de CommitJob et de la tâche peuvent différer, le chemin a été corrigé pour utiliser celui défini dans les fichiers manifestes. Pour les tâches Hive, cela signifie que le marqueur de réussite est correctement écrit lors de l'exécution

Type de modification	Description
	d'opérations telles que la partition dynamique ou UNION ALL.

Modifications apportées à Hive

Type de modification	Description
Hive est passé à la version open source 2.3.9, y compris ces correctifs JIRA	• HIVE-17155 : findConfFile () dans le HiveConf fichier .java présente quelques problèmes avec le chemin de configuration • HIVE-24797 : désactivation de la validation des valeurs par défaut lors de l'analyse des schémas Avro • HIVE-21563 : Améliorez les performances de Table# getEmptyTable en le désactivant Once registerAllFunctions • HIVE-18147 : Les tests peuvent échouer avec java.net. BindException: Adresse déjà utilisée • HIVE-24608 : revenir à get_table dans le client HMS pour Hive 2.3.x • HIVE-21200 : Vectorisation : colonne de date lançant java.lang. UnsupportedOperationException pour parquet • HIVE-19228 : suppression de l'utilisation de commons-httpclient 3.x
Rétroportages de Hive open source depuis EMR 5.34.0	• HIVE-19990 : la requête avec un intervalle littéral dans la condition de jointure échoue • HIVE-25824 : mise à niveau de la branche-2 .3 vers log4j 2.17.0

Type de modification	Description
	• TEZ-4062 : la planification des tentatives spéculatives doit être abandonnée une fois la tâche terminée • TEZ-4108 : NullPointerException lors d'une exécution spéculative, condition de course • TEZ-3918 : le paramètre tez.task.log.level ne fonctionne pas
Mises à niveau et correctifs apportés à Hive	• Mise à niveau de la version de Log4j vers 2.17.1 • Mise à niveau de la version de ORC vers 1.4.3 • Correction d'un blocage dû à un fil de pénalité ShuffleScheduler
Nouvelles fonctionnalités	• Ajout d'une fonctionnalité permettant d'imprimer Hive Query dans les journaux AM. Par défaut, l'option est désactivée. Drapeau/Conf :.`tez.am.emr.print.hive.query.in.log` Statut (par défaut) : FALSE.

Modifications apportées à Oozie

Type de modification	Description
Rétroportages d'Oozie open source depuis EMR 5.34.0	• OOZIE-3652 : Le lanceur Oozie devrait réessayer de répertorier les répertoires lorsque cela se produit NoSuchFileException

Modifications apportées à Pig

Type de modification	Description
Mises à niveau	• log4j a été mis à jour vers la version 1.2.17.

Problèmes connus

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&‘()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Version 5.34.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.34.0. Les modifications ont été apportées à la version 5.33.1.

Date de parution initiale : 20 janvier 2022

Date de mise à niveau : 21 mars 2022

Nouvelles fonctions

- [Mise à l'échelle gérée] Optimisation de la mise à l'échelle gérée des données de réorganisation Spark – Pour Amazon EMR versions 5.34.0 et ultérieures, et EMR versions 6.4.0 et ultérieures, la mise à l'échelle gérée prend désormais en compte les données de réorganisation Spark (données que Spark redistribue entre les partitions pour effectuer des opérations spécifiques). Pour plus d'informations sur les opérations de réorganisation, consultez Utilisation de la mise à l'échelle gérée par EMR dans Amazon EMR dans le Guide de gestion Amazon EMR et le Guide de programmation Spark.
- [Hudi] Améliorations visant à simplifier la configuration de Hudi. Le contrôle de simultanéité optimiste a été désactivé par défaut.

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.
- Auparavant, le redémarrage manuel du gestionnaire de ressources sur un cluster multimaître provoquait le rechargement par les démons Amazon EMR on-cluster, comme Zookeeper, de tous les nœuds précédemment mis hors service ou perdus dans le fichier znode de Zookeeper. Cela a entraîné le dépassement des limites par défaut dans certaines situations. Amazon EMR supprime désormais les enregistrements de nœuds mis hors service ou perdus datant de plus d'une heure du fichier Zookeeper et les limites internes ont été augmentées.
- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.

- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Mise à niveau de Zeppelin vers la version 0.10.0.
- Livy Fix – mise à niveau vers la version 0.7.1
- Amélioration des performances de Spark – les exécuteurs hétérogènes sont désactivés lorsque certaines valeurs de configuration de Spark sont remplacées dans EMR 5.34.0.
- WebHDFS et le serveur HttpFS sont désactivés par défaut. Vous pouvez réactiver WebHDFS en utilisant la configuration Hadoop, `dfs.webhdfs.enabled`. Le serveur HttpFS peut être démarré en utilisant `sudo systemctl start hadoop-httpfs`.

Problèmes connus

- La fonctionnalité Blocs-notes Amazon EMR utilisée avec l'emprunt d'identité de l'utilisateur Livy ne fonctionne pas car HttpFS est désactivé par défaut. Dans ce cas, le bloc-notes EMR ne peut pas se connecter au cluster dont l'emprunt d'identité Livy est activé. La solution consiste à démarrer le serveur HttpFS avant de connecter le bloc-notes EMR au cluster à l'aide de `sudo systemctl start hadoop-httpfs`.
- Les requêtes Hue ne fonctionnent pas dans Amazon EMR 6.4.0 car le serveur Apache Hadoop HTTPFS est désactivé par défaut. Pour utiliser Hue sur Amazon EMR 6.4.0, démarrez manuellement le serveur HttpFS sur le nœud primaire d'Amazon EMR à l'aide de `sudo systemctl start hadoop-httpfs`, ou utilisez une étape d'Amazon EMR.
- La fonctionnalité Blocs-notes Amazon EMR utilisée avec l'emprunt d'identité de l'utilisateur Livy ne fonctionne pas car HttpFS est désactivé par défaut. Dans ce cas, le bloc-notes EMR ne peut pas se connecter au cluster dont l'emprunt d'identité Livy est activé. La solution consiste à démarrer le serveur HttpFS avant de connecter le bloc-notes EMR au cluster à l'aide de `sudo systemctl start hadoop-httpfs`.
- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster

de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :

- Deux partitions ou plus sont analysées à partir de la même table.
- Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
- Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Version 6.5.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.5.0. Les modifications ont été apportées à la version 6.4.0.

Date de parution initiale : 20 janvier 2022

Date de mise à niveau : 21 mars 2022

Nouvelles fonctions

- [Mise à l'échelle gérée] Optimisation de la mise à l'échelle gérée des données de réorganisation Spark – Pour Amazon EMR versions 5.34.0 et ultérieures, et EMR versions 6.4.0 et ultérieures, la mise à l'échelle gérée prend désormais en compte les données de réorganisation Spark (données que Spark redistribue entre les partitions pour effectuer des opérations spécifiques). Pour plus d'informations sur les opérations de réorganisation, consultez Utilisation de la mise à l'échelle gérée par EMR dans Amazon EMR dans le Guide de gestion Amazon EMR et le Guide de programmation Spark.
- À partir d'Amazon EMR 5.32.0 et 6.5.0, le dimensionnement dynamique de l'exécuteur pour Apache Spark est activé par défaut. Pour activer ou désactiver cette fonctionnalité, vous pouvez utiliser le paramètre de configuration `spark.yarn.heterogeneousExecutors.enabled`.

- Prise en charge du format de table ouvert Apache Iceberg pour les jeux de données analytiques volumineux.
- Support pour ranger-trino-plugin 2.0.1-amzn-1
- Prise en charge de toree 0.5.0

Modifications, améliorations et problèmes résolus

- La version 6.5 d'Amazon EMR prend désormais en charge Apache Iceberg 0.12.0 et apporte des améliorations d'exécution avec l'environnement d'exécution Amazon EMR pour Apache Spark, l'environnement d'exécution Amazon EMR pour Presto et l'environnement d'exécution Amazon EMR pour Apache Hive.
- Apache Iceberg est un format de table ouvert pour les grands jeux de données dans Amazon S3. Il fournit des performances de requête rapides sur de grandes tables, des validations atomiques, des écritures simultanées et une évolution de table compatible avec SQL. Avec EMR 6.5, vous pouvez utiliser Apache Spark 3.1.2 avec le format de table Iceberg.
- Apache Hudi 0.9 ajoute la prise en charge de DDL et DML de Spark SQL. Cela vous permet de créer et de modifier des tables Hudi en utilisant uniquement des instructions SQL. Apache Hudi 0.9 inclut également des améliorations des performances côté requête et côté écriture.
- L'environnement d'exécution Amazon EMR pour Apache Hive améliore les performances d'Apache Hive sur Amazon S3 en supprimant les opérations de changement de nom pendant les opérations intermédiaires et en améliorant les performances des commandes de vérification du métastore (MSCK) utilisées pour la réparation des tables.

Problèmes connus

- Lorsque les versions 6.5.0, 6.6.0 ou 6.7.0 d'Amazon EMR lisent les tables Apache Phoenix via le shell Apache Spark, une erreur `NoSuchMethodError` se produit car Amazon EMR utilise une `Hbase.compat.version` incorrecte. La version 6.8.0 d'Amazon EMR résout ce problème.
- Les clusters de la solution groupée Hbase en haute disponibilité (HA) ne parviennent pas à se provisionner avec la taille de volume et le type d'instance par défaut. La solution à ce problème consiste à augmenter la taille du volume racine.
- Pour utiliser les actions Spark avec Apache Oozie, vous devez ajouter la configuration suivante à votre fichier Oozie `workflow.xml`. Sinon, plusieurs bibliothèques critiques telles que Hadoop et EMRFS seront absentes du classpath des exécuteurs Spark lancés par Oozie.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Version 6.4.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.4.0. Les modifications ont été apportées à la version 6.3.0.

Date de parution initiale : 20 septembre 2021

Date de mise à niveau : 21 mars 2022

Applications prises en charge

- AWS SDK pour Java version 1.12.31
- CloudWatch Version 2.2.0 de l'évier
- Connecteur DynamoDB version 4.16.0
- EMRFS version 2.47.0

- Amazon EMR Goodies version 3.2.0
- Amazon EMR Kinesis Connector version 3.5.0
- Amazon EMR Record Server version 2.1.0
- Amazon EMR Scripts version 2.5.0
- Flink version 1.13.1
- Ganglia version 3.7.2
- AWS Client Glue Hive Metastore version 3.3.0
- Hadoop version 3.2.1-amzn-4
- HBase version 2.4.4-amzn-0
- HBase-operator-tools 1.1.0
- HCatalog version 3.1.2-amzn-5
- Hive version 3.1.2-amzn-5
- Hudi version 0.8.0-amzn-0
- Hue version 4.9.0
- Java JDK version Corretto-8.302.08.1 (build 1.8.0_302-b08)
- JupyterHub version 1.4.1
- Livy version 0.7.1-incubating
- MXNet version 1.8.0
- Oozie version 5.2.1
- Phoenix version 5.1.2
- Pig version 0.17.0
- Presto version 0.254.1-amzn-0
- Trino version 359
- Apache Ranger KMS (chiffrement transparent à plusieurs maîtres) version 2.0.0
- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.2.0
- SageMaker Version 1.4.1 du SDK Spark
- Scala version 2.12.10 (machine virtuelle du serveur OpenJDK 64 bits, Java 1.8.0_282)
- Spark version 3.1.2-amzn-0

- spark-rapids 0.4.1
- Sqoop version 1.4.7
- TensorFlow version 2.4.1
- tez version 0.9.2
- Zeppelin version 0.9.0
- Zookeeper version 3.5.7
- Connecteurs et pilotes : Connecteur DynamoDB 4.16.0

Nouvelles fonctionnalités

- [Mise à l'échelle gérée] Optimisation de la mise à l'échelle gérée des données de réorganisation Spark – Pour Amazon EMR versions 5.34.0 et ultérieures, et EMR versions 6.4.0 et ultérieures, la mise à l'échelle gérée prend désormais en compte les données de réorganisation Spark (données que Spark redistribue entre les partitions pour effectuer des opérations spécifiques). Pour plus d'informations sur les opérations de réorganisation, consultez Utilisation de la mise à l'échelle gérée par EMR dans Amazon EMR dans le Guide de gestion Amazon EMR et le Guide de programmation Spark.
- Sur les clusters Amazon EMR compatibles avec Apache Ranger, vous pouvez utiliser Apache Spark SQL pour insérer des données dans les tables de métastore Apache Hive ou les mettre à jour à l'aide de `INSERT INTO`, `INSERT OVERWRITE` et `ALTER TABLE`. Lorsque vous utilisez ALTER TABLE avec Spark SQL, l'emplacement d'une partition doit être le répertoire enfant d'un emplacement de table. Amazon EMR ne prend actuellement pas en charge l'insertion de données dans une partition où l'emplacement de la partition est différent de celui de la table.
- PrestoSQL a été renommé Trino.
- Hive : l'exécution de requêtes SELECT simples avec la clause LIMIT est accélérée en arrêtant l'exécution de la requête dès que le nombre d'enregistrements mentionné dans la clause LIMIT est récupéré. Les requêtes SELECT simples sont des requêtes qui ne contiennent pas de clause GROUP BY/ORDER BY ou des requêtes qui n'ont pas d'étape de réduction. Par exemple, `SELECT * from <TABLE> WHERE <Condition> LIMIT <Number>`.

Contrôles de simultanéité Hudi

- Hudi prend désormais en charge le contrôle de simultanéité optimiste (OCC), qui peut être exploité avec des opérations d'écriture telles que UPSERT et INSERT pour permettre les modifications de plusieurs enregistreurs sur la même table Hudi. Il s'agit d'un OCC au niveau du fichier, de sorte

que deux validations (ou enregistreurs) peuvent écrire dans la même table, si leurs modifications n'entrent pas en conflit. Pour plus d'informations, consultez le contrôle de simultanéité de Hudi.

- Zookeeper est installé sur les clusters Amazon EMR, qui peut être utilisé comme fournisseur de verrous pour OCC. Pour faciliter l'utilisation de cette fonctionnalité, les propriétés préconfigurées des clusters Amazon EMR sont les suivantes :

```
hoodie.write.lock.provider=org.apache.hudi.client.transaction.lock.ZookeeperBasedLockProvider
hoodie.write.lock.zookeeper.url=<EMR Zookeeper URL>
hoodie.write.lock.zookeeper.port=<EMR Zookeeper Port>
hoodie.write.lock.zookeeper.base_path=/hudi
```

Pour activer l'OCC, vous devez configurer les propriétés suivantes soit avec leurs options de tâche Hudi, soit au niveau du cluster à l'aide de l'API de configuration Amazon EMR :

```
hoodie.write.concurrency.mode=optimistic_concurrency_control
hoodie.cleaner.policy.failed.writes=LAZY (Performs cleaning of failed writes lazily
 instead of inline with every write)
hoodie.write.lock.zookeeper.lock_key=<Key to uniquely identify the Hudi table> (Table
 Name is a good option)
```

Hudi Monitoring : CloudWatch intégration d'Amazon pour générer des rapports sur Hudi Metrics

- Amazon EMR prend en charge la publication de Hudi Metrics sur Amazon. CloudWatch Elle est activée en définissant les configurations requises suivantes :

```
hoodie.metrics.on=true
hoodie.metrics.reporter.type=CLOUDWATCH
```

- Les configurations Hudi facultatives que vous pouvez modifier sont les suivantes :

Paramètre	Description	Valeur
hoodie.metrics.cloudwatch.report.period.seconds	Fréquence (en secondes) à laquelle les statistiques doivent être communiquées à Amazon CloudWatch	La valeur par défaut est de 60 s, ce qui convient à la résolution d'une minute par défaut proposée par Amazon CloudWatch

Paramètre	Description	Valeur
hoodie.metrics.cloudwatch.metric.prefix	Préfixe à ajouter au nom de chaque métrique	La valeur par défaut est vide (pas de préfixe)
hoodie.metrics.cloudwatch.namespace	CloudWatch Espace de noms Amazon sous lequel les métriques sont publiées	La valeur par défaut est Hudi
hoodie.metrics.cloudwatch.maxDatumsPerDemande	Nombre maximum de références à inclure dans une demande adressée à Amazon CloudWatch	La valeur par défaut est 20, identique à la valeur CloudWatch par défaut d'Amazon

Prise en charge et amélioration des configurations Hudi d'Amazon EMR

- Les clients peuvent désormais tirer parti de l'API de configuration EMR et de la fonctionnalité de reconfiguration pour configurer les configurations Hudi au niveau du cluster. Un nouveau support de configuration basé sur des fichiers a été introduit via/etc/hudi/conf/hudi-defaults.conf, à l'instar d'autres applications telles que Spark, Hive, etc. EMR configure quelques paramètres par défaut pour améliorer l'expérience utilisateur :

  — `hoodie.datasource.hive_sync.jdbcurl` est configuré selon l'URL du serveur Hive du cluster et n'a plus besoin d'être spécifié. Cela est particulièrement utile lorsque vous exécutez une tâche en mode cluster Spark, où vous deviez auparavant spécifier l'adresse IP principale Amazon EMR.

  — HBase des configurations spécifiques, utiles pour utiliser l' HBase index avec Hudi.

  — Configuration spécifique au fournisseur de verrous Zookeeper, comme indiqué dans la section Contrôle de simultanéité, qui facilite l'utilisation du contrôle de simultanéité optimiste (OCC).
- Des modifications supplémentaires ont été introduites pour réduire le nombre de configurations à transmettre et pour en déduire automatiquement dans la mesure du possible :

  — Le mot clé `partitionBy` peut être utilisé pour spécifier la colonne de partition.

— Lorsque vous activez Hive Sync, il n'est plus obligatoire de spécifier `HIVE_TABLE_OPT_KEY, HIVE_PARTITION_FIELDS_OPT_KEY, HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY`. Ces valeurs peuvent être déduites du nom de la table Hudi et du champ de partition.

— `KEYGENERATOR_CLASS_OPT_KEY` n'est pas obligatoire et peut être déduite de cas plus simples de `SimpleKeyGenerator` et`ComplexKeyGenerator`.

Mises en garde de Hudi

- Hudi ne prend pas en charge l'exécution vectorisée dans les tables Hive for Merge on Read (MoR) et Bootstrap. Par exemple, `count(*)` échoue avec la table en temps réel de Hudi lorsque `hive.vectorized.execution.enabled` est défini sur true. Comme solution de contournement, vous pouvez désactiver la lecture vectorisée en définissant `hive.vectorized.execution.enabled` sur `false`.
- La prise en charge des enregistreurs multiples n'est pas compatible avec la fonction d'amorçage de Hudi.
- Flink Streamer et Flink SQL sont des fonctionnalités expérimentales dans cette version. Ces fonctionnalités ne sont pas recommandées pour les déploiements de production.

Modifications, améliorations et problèmes résolus

Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.

- Auparavant, le redémarrage manuel du gestionnaire de ressources sur un cluster multimaître provoquait le rechargement par les démons Amazon EMR on-cluster, comme Zookeeper, de tous les nœuds précédemment mis hors service ou perdus dans le fichier znode de Zookeeper. Cela a entraîné le dépassement des limites par défaut dans certaines situations. Amazon EMR supprime désormais les enregistrements de nœuds mis hors service ou perdus datant de plus d'une heure du fichier Zookeeper et les limites internes ont été augmentées.
- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.

- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Configuration d'un cluster pour résoudre les problèmes de performances d'Apache YARN Timeline Server versions 1 et 1.5

  Les versions 1 et 1.5 d'Apache YARN Timeline Server peuvent entraîner des problèmes de performances avec de grands clusters EMR très actifs, en particulier avec `yarn.resourcemanager.system-metrics-publisher.enabled=true`, le paramètre par défaut d'Amazon EMR. YARN Timeline Server v2 open source résout le problème de performance lié à la capacité de mise à l'échelle de YARN Timeline Server.

  Les autres solutions à ce problème incluent :

  - Configuration de yarn.resourcemanager. system-metrics-publisher.enabled=false dans le fichier yarn-site.xml.
  - Activation du correctif pour ce problème lors de la création d'un cluster, comme décrit ci-dessous.

  Les versions Amazon EMR suivantes contiennent un correctif pour ce problème de performance de YARN Timeline Server.

  EMR 5.30.2, 5.31.1, 5.32.1, 5.33.1, 5.34.x, 6.0.1, 6.1.1, 6.2.1, 6.3.1, 6.4.x

  Pour activer le correctif sur l'une des versions Amazon EMR spécifiées ci-dessus, définissez ces propriétés sur `true` dans un fichier JSON de configuration transmis à l'aide du

paramètre de commande `aws emr create-cluster` : `--configurations file://./configurations.json`. Vous pouvez également activer le correctif à l'aide de l'interface utilisateur de la console de reconfiguration.

Exemple du contenu du fichier configurations.json :

```
[
{
"Classification": "yarn-site",
"Properties": {
"yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
 "true",
"yarn.resourcemanager.system-metrics-publisher.enabled": "true"
},
"Configurations": []
}
]
```

- WebHDFS et le serveur HttpFS sont désactivés par défaut. Vous pouvez réactiver WebHDFS en utilisant la configuration Hadoop, `dfs.webhdfs.enabled`. Le serveur HttpFS peut être démarré en utilisant `sudo systemctl start hadoop-httpfs`.
- Le protocole HTTPS est désormais activé par défaut pour les référentiels Amazon Linux. Si vous utilisez une politique VPCE Amazon S3 pour restreindre l'accès à des compartiments spécifiques, vous devez ajouter le nouvel ARN du compartiment Amazon Linux `arn:aws:s3:::amazonlinux-2-repos-$region/*` à votre politique (remplacez `$region` par la région où se trouve le point de terminaison). Pour plus d'informations, consultez cette rubrique dans les forums de AWS discussion. Annonce : Amazon Linux 2 permet désormais d'utiliser le protocole HTTPS lors de la connexion aux référentiels de packages.
- Hive : les performances des requêtes d'écriture sont améliorées en permettant l'utilisation d'un répertoire temporaire sur HDFS pour la dernière tâche. Les données temporaires pour la tâche finale sont écrites sur HDFS au lieu d'Amazon S3 et les performances sont améliorées car les données sont déplacées de HDFS vers l'emplacement de la table finale (Amazon S3) au lieu d'être déplacées entre les appareils Amazon S3.
- Hive : amélioration du temps de compilation des requêtes jusqu'à 2,5 fois avec l'élimination des partitions du métastore Glue.
- Par défaut, lorsque les éléments intégrés UDFs sont transmis par Hive au serveur Hive Metastore, seul un sous-ensemble de ces éléments intégrés UDFs est transmis au Glue Metastore, car Glue ne prend en charge que des opérateurs d'expression limités. Si vous définissez

`hive.glue.partition.pruning.client=true`, tout l'élimination des partitions se fait du côté client. Si vous définissez `hive.glue.partition.pruning.server=true`, tout l'élimination des partitions se fait du côté serveur.

Problèmes connus

- Les requêtes Hue ne fonctionnent pas dans Amazon EMR 6.4.0 car le serveur Apache Hadoop HTTPFS est désactivé par défaut. Pour utiliser Hue sur Amazon EMR 6.4.0, démarrez manuellement le serveur HttpFS sur le nœud primaire d'Amazon EMR à l'aide de `sudo systemctl start hadoop-httpfs`, ou utilisez une étape d'Amazon EMR.
- La fonctionnalité Blocs-notes Amazon EMR utilisée avec l'emprunt d'identité de l'utilisateur Livy ne fonctionne pas car HttpFS est désactivé par défaut. Dans ce cas, le bloc-notes EMR ne peut pas se connecter au cluster dont l'emprunt d'identité Livy est activé. La solution consiste à démarrer le serveur HttpFS avant de connecter le bloc-notes EMR au cluster à l'aide de `sudo systemctl start hadoop-httpfs`.
- Dans la version 6.4.0 d'Amazon EMR, Phoenix ne prend pas en charge le composant des connecteurs Phoenix.
- Pour utiliser les actions Spark avec Apache Oozie, vous devez ajouter la configuration suivante à votre fichier Oozie `workflow.xml`. Sinon, plusieurs bibliothèques critiques telles que Hadoop et EMRFS seront absentes du classpath des exécuteurs Spark lancés par Oozie.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14

autres caractères de non-contrôle : `!"#$%&‘()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Version 5.32.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.32.0. Les modifications ont été apportées à la version 5.31.0.

Date de parution initiale : 8 janvier 2021

Mises à niveau

- Mise à niveau du connecteur Amazon Glue vers la version 1.14.0
- Mise à niveau du SDK Amazon SageMaker Spark vers la version 1.4.1
- Mise à niveau AWS SDK pour Java vers la version 1.11.890
- Mise à niveau de la version 4.16.0 du connecteur DynamoDB d'EMR
- Mise à niveau d'EMRFS vers la version 2.45.0
- Mise à niveau des métriques d'analyse des journaux d'EMR vers la version 1.18.0
- MetricsAndEventsApiGateway Client EMR mis à niveau vers la version 1.5.0
- Mise à niveau du serveur d'enregistrement EMR vers la version 1.8.0
- Mise à niveau d'EMR S3 Dist CP vers la version 2.17.0
- Mise à niveau d'EMR Secret Agent vers la version 1.7.0
- Mise à niveau de Flink vers la version 1.11.2
- Mise à niveau de Hadoop vers la version 2.10.1-amzn-0
- Mise à niveau de Hive vers la version 2.3.7-amzn-3
- Mise à niveau de Hue vers la version 4.8.0
- Mise à niveau de Mxnet vers la version 1.7.0
- Mise à niveau d'OpenCV vers la version 4.4.0
- Mise à niveau de Presto vers la version 0.240.1-amzn-0
- Mise à niveau de Spark vers la version 2.4.7-amzn-0

- Mise à niveau TensorFlow vers la version 2.3.1

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.
- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Versions de composants mises à niveau.
- Pour obtenir la liste des versions des composants, consultez la section À propos des versions d'Amazon EMR dans ce guide.

Nouvelles fonctionnalités

- À partir d'Amazon EMR 5.32.0 et 6.5.0, le dimensionnement dynamique de l'exécuteur pour Apache Spark est activé par défaut. Pour activer ou désactiver cette fonctionnalité, vous pouvez utiliser le paramètre de configuration `spark.yarn.heterogeneousExecutors.enabled`.
- État de prise en charge du service de métadonnées d'instance (IMDS) V2 : les composants Amazon EMR 5.23.1, 5.27.1 et 5.32 ou versions ultérieures sont utilisés pour tous les appels IMDS. IMDSv2 Pour les appels IMDS dans le code de votre application, vous pouvez utiliser les deux IMDSv1 ou configurer l'IMDS pour qu'il ne soit utilisé que IMDSv2 pour renforcer la sécurité. IMDSv2 Pour les autres versions 5.x EMR, la IMDSv1 désactivation entraîne l'échec du démarrage du cluster.
- À partir d'Amazon EMR 5.32.0, vous pouvez lancer un cluster qui s'intègre nativement à Apache Ranger. Apache Ranger est un cadre open source permettant d'activer, de surveiller et de gérer la sécurité globale des données sur la plateforme Hadoop. Pour plus d'informations, consultez Apache Ranger. Grâce à l'intégration native, vous pouvez utiliser votre propre Apache Ranger pour appliquer un contrôle précis de l'accès aux données sur Amazon EMR. Consultez Intégration d'Amazon EMR à Apache Ranger dans le Guide de version Amazon EMR.
- La version 5.32.0 d'Amazon EMR prend en charge Amazon EMR sur EKS. Pour en savoir plus sur la prise en main d'EMR sur EKS, consultez Qu'est-ce qu'Amazon EMR sur EKS ?.
- La version 5.32.0 d'Amazon EMR prend en charge Amazon EMR Studio (version préliminaire). Pour plus d'informations sur la prise en main d'EMR Studio, consultez Amazon EMR Studio (version préliminaire).
- Politiques gérées délimitées : pour s'aligner sur les AWS meilleures pratiques, Amazon EMR a introduit des politiques gérées par défaut définies dans la version 2 EMR en remplacement des politiques qui seront déconseillées. Consultez Politiques gérées par Amazon EMR.

Problèmes connus

- Pour les clusters de sous-réseaux privés Amazon EMR 6.3.0 et 6.2.0, vous ne pouvez pas accéder à l'interface utilisateur Web de Ganglia. Vous recevrez un message d'erreur « accès refusé (403) ». D'autres sites Web UIs, tels que Spark, Hue JupyterHub, Zeppelin, Livy et Tez, fonctionnent normalement. L'accès à l'interface utilisateur Web de Ganglia sur les clusters de sous-réseaux publics fonctionne également normalement. Pour résoudre ce problème, redémarrez le service httpd sur le nœud primaire avec `sudo systemctl restart httpd`. Ce problème est résolu dans Amazon EMR 6.4.0.

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Redémarrer InstanceController

     ```
     $ sudo systemctl daemon-reload

     $ sudo systemctl restart instance-controller
     ```

  Définissez un ulimit à l'aide de l'action d'amorçage (BA)

  Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Les clusters EMR qui exécutent Amazon Linux ou Amazon Linux 2 Amazon Machine Images (AMIs) utilisent le comportement par défaut d'Amazon Linux et ne téléchargent ni n'installent automatiquement les mises à jour importantes et critiques du noyau nécessitant un redémarrage. Ce comportement est identique à celui des autres EC2 instances Amazon qui exécutent l'AMI Amazon Linux par défaut. Si de nouvelles mises à jour logicielles Amazon Linux nécessitant un redémarrage (telles que les mises à jour du noyau, de NVIDIA et de CUDA) sont disponibles après la publication d'une version d'Amazon EMR, les instances de cluster EMR qui exécutent l'AMI par défaut ne téléchargent pas et n'installent pas automatiquement ces mises à jour. Pour obtenir les mises à jour du noyau, vous pouvez personnaliser votre AMI Amazon EMR afin d'utiliser la dernière AMI Amazon Linux.

- La prise en charge par console pour créer une configuration de sécurité spécifiant l'option d'intégration de AWS Ranger n'est actuellement pas prise en charge dans la GovCloud région. La configuration de la sécurité peut être effectuée à l'aide de la CLI. Consultez Création de la configuration de sécurité EMR dans le Guide de gestion Amazon EMR.
- Lorsque AtRestEncryption le chiffrement HDFS est activé sur un cluster qui utilise Amazon EMR 5.31.0 ou 5.32.0, les requêtes Hive génèrent l'exception d'exécution suivante.

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
 attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
 java.lang.RuntimeException: Hive Runtime Error while closing
 operators: java.io.IOException: java.util.ServiceConfigurationError:
 org.apache.hadoop.security.token.TokenIdentifier: Provider
 org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Version 6.2.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.2.0. Les modifications ont été apportées à la version 6.1.0.

Date de parution initiale : 9 décembre 2020

Dernière mise à jour : 4 octobre 2021

Applications prises en charge

- AWS SDK pour Java version 1.11.828

- emr-record-server version 1.7.0
- Flink version 1.11.2
- Ganglia version 3.7.2
- Hadoop version 3.2.1-amzn-1
- HBase version 2.2.6-amzn-0
- HBase-operator-tools 1.0.0
- HCatalog version 3.1.2-amzn-0
- Hive version 3.1.2-amzn-3
- Hudi version 0.6.0-amzn-1
- Hue version 4.8.0
- JupyterHub version 1.1.0
- Livy version 0.7.0
- MXNet version 1.7.0
- Oozie version 5.2.0
- Phoenix version 5.0.0
- Pig version 0.17.0
- Presto version 0.238.3-amzn-1
- PrestoSQL version 343
- Spark version 3.0.1-amzn-0
- spark-rapids 0.2.0
- TensorFlow version 2.3.1
- Zeppelin version 0.9.0-preview1
- Zookeeper version 3.4.14
- Connecteurs et pilotes : Connecteur DynamoDB 4.16.0

Nouvelles fonctionnalités

- HBase: Suppression du changement de nom lors de la phase de validation et ajout d'un HFile suivi persistant. Consultez la section HFile Suivi permanent dans le guide de mise à jour d'Amazon EMR.

- HBase: Rétroporté Créez une configuration qui oblige à mettre en cache les blocs lors du compactage.
- PrestoDB : améliorations apportées à l'élimination dynamique des partitions. Join Reorder basée sur des règles fonctionne sur des données non partitionnées.
- Politiques gérées délimitées : pour s'aligner sur les AWS meilleures pratiques, Amazon EMR a introduit des politiques gérées par défaut définies dans la version 2 EMR en remplacement des politiques qui seront déconseillées. Consultez Politiques gérées par Amazon EMR.
- État de prise en charge du service de métadonnées d'instance (IMDS) V2 : pour Amazon EMR 6.2 ou version ultérieure, les composants Amazon EMR sont IMDSv2 utilisés pour tous les appels IMDS. Pour les appels IMDS dans le code de votre application, vous pouvez utiliser les deux IMDSv1 ou configurer l'IMDS pour qu'il ne soit utilisé que IMDSv2 pour renforcer la sécurité. IMDSv2 Si vous la désactivez IMDSv1 dans les versions antérieures d'Amazon EMR 6.x, cela entraîne un échec du démarrage du cluster.

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.
- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification

Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.

- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Spark : amélioration des performances dans l'environnement d'exécution de Spark.

Problèmes connus

- Les autorisations définies sur Amazon EMR 6.2 sont incorrectes. Le répertoire/etc/cron.d/libinstance-controller-java file in EMR 6.2.0. Permissions on the file are 645 (-rw-r--r-x), when they should be 644 (-rw-r--r--). As a result, Amazon EMR version 6.2 does not log instance-state logs, and the /emr/instance-logs est vide. Ce problème est corrigé dans Amazon EMR 6.3.0 et les versions ultérieures.

  Pour contourner ce problème, exécutez le script suivant en tant qu'action d'amorçage lors du lancement du cluster.

  ```
  #!/bin/bash
  sudo chmod 644 /etc/cron.d/libinstance-controller-java
  ```

- Pour les clusters de sous-réseaux privés Amazon EMR 6.2.0 et 6.3.0, vous ne pouvez pas accéder à l'interface utilisateur Web de Ganglia. Vous recevrez un message d'erreur « accès refusé (403) ». D'autres sites Web UIs, tels que Spark, Hue JupyterHub, Zeppelin, Livy et Tez, fonctionnent normalement. L'accès à l'interface utilisateur Web de Ganglia sur les clusters de sous-réseaux publics fonctionne également normalement. Pour résoudre ce problème, redémarrez le service httpd sur le nœud primaire avec `sudo systemctl restart httpd`. Ce problème est résolu dans Amazon EMR 6.4.0.
- Amazon EMR 6.2.0 présente un problème selon lequel httpd échoue continuellement, ce qui rend Ganglia indisponible. Le message d'erreur « Impossible de se connecter au serveur » s'affiche. Pour réparer un cluster déjà en cours d'exécution présentant ce problème, connectez-vous en SSH au nœud primaire du cluster et ajoutez la ligne `Listen 80` au fichier `httpd.conf` situé dans `/etc/httpd/conf/httpd.conf`. Ce problème est résolu dans Amazon EMR 6.3.0.

- HTTTD échoue sur les clusters EMR 6.2.0 lorsque vous utilisez une configuration de sécurité. Cela rend l'interface utilisateur de l'application web Ganglia indisponible. Pour accéder à l'interface utilisateur de l'application web Ganglia, ajoutez `Listen 80` au fichier `/etc/httpd/conf/httpd.conf` sur le nœud primaire de votre cluster. Pour plus d'informations sur la connexion à votre cluster, consultez Connexion au nœud primaire à l'aide de SSH.

  Blocs-notes EMR ne parvient pas non plus à établir une connexion avec les clusters EMR 6.2.0 lorsque vous utilisez une configuration de sécurité. Le bloc-notes ne parviendra pas à répertorier les noyaux et à soumettre les tâches Spark. Nous vous recommandons d'utiliser Blocs-notes EMR avec une autre version d'Amazon EMR à la place.

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

2. Redémarrer InstanceController

```
$ sudo systemctl daemon-reload

$ sudo systemctl restart instance-controller
```

Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Amazon EMR 6.1.0 et 6.2.0 présentent un problème de performance qui peut affecter de manière critique toutes les opérations insert, upsert et delete de Hudi. Si vous envisagez d'utiliser Hudi avec Amazon EMR 6.1.0 ou 6.2.0, AWS contactez le support pour obtenir un RPM Hudi corrigé.

- **Important**

  Les clusters EMR qui exécutent Amazon Linux ou Amazon Linux 2 Amazon Machine Images (AMIs) utilisent le comportement par défaut d'Amazon Linux et ne téléchargent ni n'installent automatiquement les mises à jour importantes et critiques du noyau nécessitant un redémarrage. Ce comportement est identique à celui des autres EC2 instances Amazon qui exécutent l'AMI Amazon Linux par défaut. Si de nouvelles mises à jour logicielles Amazon Linux nécessitant un redémarrage (telles que les mises à jour du noyau, de NVIDIA et de CUDA) sont disponibles après la publication d'une version d'Amazon EMR, les instances de cluster EMR qui exécutent l'AMI par défaut ne téléchargent pas et n'installent pas automatiquement ces mises à jour. Pour obtenir les mises à jour du noyau, vous pouvez personnaliser votre AMI Amazon EMR afin d'utiliser la dernière AMI Amazon Linux.

- Les artefacts Maven d'Amazon EMR 6.2.0 ne sont pas publiés. Ils seront publiés avec une future version d'Amazon EMR.
- Le HFile suivi permanent à l'aide de la table système HBase Storefile ne prend pas en charge la fonctionnalité de réplication HBase régionale. Pour plus d'informations sur HBase la réplication régionale, consultez la section Nombre élevé de lectures disponibles cohérentes avec la chronologie.
- Différences de version entre Amazon EMR 6.x et EMR 5.x pour la compartimentation Hive

  EMR 5.x utilise OOS Apache Hive 2, tandis que EMR 6.x utilise OOS Apache Hive 3. La version open source Hive2 utilise la version 1 de Bucketing, tandis que la version open source Hive3 utilise la version 2. Cette différence de version de compartimentation entre Hive 2 (EMR 5.x) et Hive 3 (EMR 6.x) signifie que le hachage de compartimentation de Hive fonctionne différemment. Consultez l'exemple ci-dessous.

  Le tableau suivant est un exemple créé dans EMR 6.x et EMR 5.x, respectivement.

```
-- Using following LOCATION in EMR 6.x
CREATE TABLE test_bucketing (id INT, desc STRING)
PARTITIONED BY (day STRING)
CLUSTERED BY(id) INTO 128 BUCKETS
LOCATION 's3://your-own-s3-bucket/emr-6-bucketing/';

-- Using following LOCATION in EMR 5.x
LOCATION 's3://your-own-s3-bucket/emr-5-bucketing/';
```

Insertion des mêmes données dans EMR 6.x et EMR 5.x.

```
INSERT INTO test_bucketing PARTITION (day='01') VALUES(66, 'some_data');
INSERT INTO test_bucketing PARTITION (day='01') VALUES(200, 'some_data');
```

La vérification de l'emplacement S3 montre que le nom du fichier de compartimentation est différent, car la fonction de hachage est différente entre EMR 6.x (Hive 3) et EMR 5.x (Hive 2).

```
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-6-bucketing/day=01/
2020-10-21 20:35:16         13 000025_0
2020-10-21 20:35:22         14 000121_0
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-5-bucketing/day=01/
2020-10-21 20:32:07         13 000066_0
2020-10-21 20:32:51         14 000072_0
```

Vous pouvez également constater la différence de version en exécutant la commande suivante dans la CLI Hive dans EMR 6.x. Notez qu'il renvoie la version 2 de compartimentation.

```
hive> DESCRIBE FORMATTED test_bucketing;
...
Table Parameters:
    bucketing_version       2
...
```

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur hadoop au nœud primaire du cluster EMR avec plusieurs nœuds primaires.

- Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

  ```
  kinit -kt <keytab_file> <principal>
  ```

  Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Version 5.31.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.31.0. Les modifications ont été apportées à la version 5.30.1.

Date de parution initiale : 9 octobre 2020

Dernière mise à jour : 15 octobre 2020

Mises à niveau

- Mise à niveau du connecteur Amazon Glue vers la version 1.13.0
- Mise à niveau du SDK Amazon SageMaker Spark vers la version 1.4.0
- Mise à niveau du connecteur Amazon Kinesis vers la version 3.5.9
- Mise à niveau AWS SDK pour Java vers la version 1.11.852
- Mise à niveau de Bigtop-tomcat vers la version 8.5.56
- Mise à niveau d'EMRFS vers la version 2.43.0
- MetricsAndEventsApiGateway Client EMR mis à niveau vers la version 1.4.0
- Mise à niveau d'EMR S3 Dist CP vers la version 2.15.0
- Mise à niveau d'EMR S3 Select vers la version 1.6.0
- Mise à niveau de Flink vers la version 1.11.0
- Mise à niveau de Hadoop vers la version 2.10.0
- Mise à niveau de Hive vers la version 2.3.7
- Mise à niveau de Hudi vers la version 0.6.0
- Mise à niveau de Hue vers la version 4.7.1
- Mise à niveau JupyterHub vers la version 1.1.0
- Mise à niveau de Mxnet vers la version 1.6.0
- Mise à niveau d'OpenCV vers la version 4.3.0
- Mise à niveau de Presto vers la version 0.238.3
- Mise à niveau TensorFlow vers la version 2.1.0

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.

- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Les statistiques des colonnes Hive sont prises en charge pour les versions 5.31.0 et ultérieures d'Amazon EMR.
- Versions de composants mises à niveau.
- Prise en charge d'EMRFS S3EC V2 dans Amazon EMR 5.31.0. Dans les versions 1.11.837 et ultérieures du kit SDK Java S3, le client de chiffrement version 2 (S3EC V2) a été introduit avec diverses améliorations de sécurité. Pour plus d'informations, consultez les ressources suivantes :
  - Article de blog S3 : Updates to the Amazon S3 encryption client.
  - AWS SDK pour Java Guide du développeur : Migrez les clients de chiffrement et de déchiffrement vers la version V2.

- Guide de gestion EMR : Chiffrement côté client Amazon S3.

Le client de chiffrement V1 est toujours disponible dans le kit SDK pour des raisons de rétrocompatibilité.

Nouvelles fonctionnalités

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Redémarrer InstanceController

     ```
     $ sudo systemctl daemon-reload
     ```

```
$ sudo systemctl restart instance-controller
```

Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Avec Amazon EMR 5.31.0, vous pouvez lancer un cluster qui s'intègre à Lake Formation. Cette intégration fournit un filtrage des données précis au niveau des colonnes pour les bases de données et les tables du Glue AWS Data Catalog. Il permet également une authentification unique fédérée à Blocs-notes EMR ou Apache Zeppelin à partir d'un système d'identité d'entreprise. Pour plus d'informations, consultez Intégration d'Amazon EMR avec AWS Lake Formation dans le Guide de gestion Amazon EMR.

  Amazon EMR with Lake Formation est actuellement disponible dans 16 AWS régions : USA Est (Ohio et Virginie du Nord), USA Ouest (Californie du Nord et Oregon), Asie-Pacifique (Mumbai, Séoul, Singapour, Sydney et Tokyo), Canada (Centre), Europe (Francfort, Irlande, Londres, Paris et Stockholm), Amérique du Sud (São Paulo).

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

- Lorsque AtRestEncryption le chiffrement HDFS est activé sur un cluster qui utilise Amazon EMR 5.31.0 ou 5.32.0, les requêtes Hive génèrent l'exception d'exécution suivante.

  ```
  TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
   attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
   java.lang.RuntimeException: Hive Runtime Error while closing
   operators: java.io.IOException: java.util.ServiceConfigurationError:
  ```

```
org.apache.hadoop.security.token.TokenIdentifier: Provider
org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère / (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Version 6.1.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.1.0. Les modifications ont été apportées à la version 6.0.0.

Date de parution initiale : 4 septembre 2020

Dernière mise à jour : 15 octobre 2020

Applications prises en charge

- AWS SDK pour Java version 1.11.828
- Flink version 1.11.0
- Ganglia version 3.7.2
- Hadoop version 3.2.1-amzn-1
- HBase version 2.2.5

- HBase-operator-tools 1.0.0
- HCatalog version 3.1.2-amzn-0
- Hive version 3.1.2-amzn-1
- Hudi version 0.5.2-incubation
- Hue version 4.7.1
- JupyterHub version 1.1.0
- Livy version 0.7.0
- MXNet version 1.6.0
- Oozie version 5.2.0
- Phoenix version 5.0.0
- Presto version 0.232
- PrestoSQL version 338
- Spark version 3.0.0-amzn-0
- TensorFlow version 2.1.0
- Zeppelin version 0.9.0-preview1
- Zookeeper version 3.4.14
- Connecteurs et pilotes : Connecteur DynamoDB 4.14.0

Nouvelles fonctionnalités

- Les types d'instance ARM sont pris en charge à partir des versions 5.30.0 et 6.1.0 d'Amazon EMR.
- Les types d'instances à usage général M6g sont pris en charge à partir des versions 6.1.0 et 5.30.0 d'Amazon EMR. Pour plus d'informations, consultez Types d'instance pris en charge dans le Guide de gestion Amazon EMR.
- La fonctionnalité de groupe de EC2 placement est prise en charge à partir de la version 5.23.0 d'Amazon EMR en tant qu'option pour plusieurs clusters de nœuds principaux. Actuellement, seuls les types de nœuds primaires sont pris en charge par la fonctionnalité de groupe de placement. La stratégie `SPREAD` est appliquée à ces nœuds primaires. La stratégie `SPREAD` place un petit groupe d'instances sur un matériel sous-jacent distinct afin de se prémunir contre la perte de plusieurs nœuds primaires en cas de panne matérielle. Pour plus d'informations, consultez la section Intégration de l'EMR au groupe de EC2 placement dans le guide de gestion Amazon EMR.
- Mise à l'échelle gérée – Avec la version 6.1.0 d'Amazon EMR, vous pouvez activer la mise à l'échelle gérée par Amazon EMR pour augmenter ou diminuer automatiquement le nombre

d'instances ou d'unités dans votre cluster en fonction de la charge de travail. Amazon EMR évalue en permanence les métriques de cluster pour prendre des décisions de dimensionnement qui optimisent vos clusters en termes de coût et de vitesse. Managed Scaling est également disponible sur Amazon EMR version 5.30.0 et suivantes, à l'exception de 6.0.0. Pour plus d'informations, consultez Mise à l'échelle des ressources de cluster dans le Guide de gestion Amazon EMR.

- La version 338 de PrestoSQL est prise en charge avec EMR 6.1.0. Pour plus d'informations, consultez Presto.
  - PrestoSQL est uniquement pris en charge sur EMR 6.1.0 et versions ultérieures, pas sur EMR 6.0.0 ou EMR 5.x.
  - Le nom de l'application, `Presto` continue d'être utilisé pour installer PrestoDB sur des clusters. Pour installer PrestoSQL sur des clusters, utilisez le nom de l'application `PrestoSQL`.
  - Vous pouvez installer PrestoDB ou PrestoSQL, mais vous ne pouvez pas installer les deux sur un seul cluster. Si PrestoDB et PrestoSQL sont spécifiés lors de la tentative de création d'un cluster, une erreur de validation se produit et la demande de création de cluster échoue.
  - PrestoSQL est pris en charge sur les clusters mono-maître et muti-maître. Sur les clusters multi-maîtres, un métastore Hive externe est requis pour exécuter PrestoSQL ou Prestodb. Consultez Applications prises en charge dans un cluster EMR avec plusieurs nœuds primaires.
- Prise en charge de l'authentification automatique ECR sur Apache Hadoop et Apache Spark avec Docker : les utilisateurs de Spark peuvent utiliser des images Docker provenant de Docker Hub et d'Amazon Elastic Container Registry (Amazon ECR) pour définir les dépendances de l'environnement et des bibliothèques.

  Configurez Docker et exécutez des applications Spark avec Docker à l'aide d'Amazon EMR 6.x.
- EMR prend en charge les transactions ACID d'Apache Hive : Amazon EMR 6.1.0 prend en charge les transactions ACID de Hive afin qu'elles soient conformes aux propriétés ACID d'une base de données. Grâce à cette fonctionnalité, vous pouvez exécuter les opérations `INSERT`, `UPDATE`, `DELETE`, et `MERGE` dans des tables gérées par Hive avec des données stockées dans Amazon Simple Storage Service (Amazon S3). Il s'agit d'une fonctionnalité essentielle pour les cas d'utilisation tels que l'ingestion en continu, le retraitement des données, les mises à jour en masse à l'aide de MERGE et les dimensions qui changent lentement. Pour plus d'informations, notamment des exemples de configuration et des cas supports, consultez Amazon EMR prend en charge les transactions ACID d'Apache Hive.

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.
- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.
- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Apache Flink n'est pas pris en charge sur EMR 6.0.0, mais il l'est sur EMR 6.1.0 avec Flink 1.11.0. Il s'agit de la première version de Flink officiellement compatible avec Hadoop 3. Consultez Annonce de sortie d'Apache Flink 1.11.0.
- Ganglia a été supprimé des packages EMR 6.1.0 par défaut.

Problèmes connus

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Redémarrer InstanceController

     ```
     $ sudo systemctl daemon-reload

     $ sudo systemctl restart instance-controller
     ```

  Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Amazon EMR 6.1.0 et 6.2.0 présentent un problème de performance qui peut affecter de manière critique toutes les opérations insert, upsert et delete de Hudi. Si vous envisagez d'utiliser Hudi avec Amazon EMR 6.1.0 ou 6.2.0, AWS contactez le support pour obtenir un RPM Hudi corrigé.

- Si vous définissez une configuration de collecte des déchets personnalisée avec `spark.driver.extraJavaOptions` et`spark.executor.extraJavaOptions`, cela entraînera un échec du driver/executor lancement d'EMR 6.1 en raison d'une configuration de collecte des déchets conflictuelle. Avec la version 6.1.0 d'EMR, vous devez spécifier la configuration personnalisée du récupérateur de mémoire Spark pour les pilotes et les exécuteurs avec les propriétés `spark.driver.defaultJavaOptions` et `spark.executor.defaultJavaOptions` à la place. Pour en savoir plus, consultez Environnement d'exécution Apache Spark et Configuration du récupérateur de mémoire Spark sur Amazon EMR 6.1.0.

- L'utilisation de Pig avec Oozie (et au sein de Hue, puisque Hue utilise les actions Oozie pour exécuter les scripts Pig), génère une erreur indiquant qu'une bibliothèque native-lzo ne peut pas être chargée. Ce message d'erreur est informatif et n'empêche pas Pig de fonctionner.
- Prise en charge de la simultanéité dans Hudi : Actuellement, Hudi ne prend pas en charge les écritures simultanées dans une seule table Hudi. De plus, Hudi annule toutes les modifications effectuées par les enregistreurs en cours avant de permettre à un nouvel enregistreur de démarrer. Les écritures simultanées peuvent interférer avec ce mécanisme et introduire des conditions de concurrence, ce qui peut entraîner une corruption des données. Vous devez vous assurer que, dans le cadre de votre flux de traitement des données, il n'y a qu'un seul enregistreur Hudi opérant sur une table Hudi à tout moment. Hudi prend en charge plusieurs lecteurs simultanés opérant sur la même table Hudi.
- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos

paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

- Il y a un problème dans Amazon EMR 6.1.0 qui affecte les clusters exécutant Presto. Après une période prolongée (plusieurs jours), le cluster peut générer des erreurs telles que "su : failed to execute /bin/bash : Resource temporarily unavailable" ou "shell request failed on channel 0". Ce problème est dû à un processus interne Amazon EMR (InstanceController) qui génère trop de processus légers (LWP), ce qui finit par amener l'utilisateur Hadoop à dépasser sa limite nproc. Cela empêche l'utilisateur d'ouvrir des processus supplémentaires. La solution à ce problème est de mettre à niveau vers EMR 6.2.0.

Version 6.0.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 6.0.0.

Date de parution initiale : 10 mars 2020

Applications prises en charge

- AWS SDK pour Java version 1.11.711
- Ganglia version 3.7.2
- Hadoop version 3.2.1
- HBase version 2.2.3
- HCatalog version 3.1.2
- Hive version 3.1.2
- Hudi version 0.5.0-incubation
- Hue version 4.4.0
- JupyterHub version 1.0.0
- Livy version 0.6.0
- MXNet version 1.5.1
- Oozie version 5.1.0
- Phoenix version 5.0.0
- Presto version 0.230
- Spark version 2.4.4

- TensorFlow version 1.14.0
- Zeppelin version 0.9.0-SNAPSHOT
- Zookeeper version 3.4.14
- Connecteurs et pilotes : Connecteur DynamoDB 4.14.0

> **Note**
>
> Flink, Sqoop, Pig et Mahout ne sont pas disponibles dans Amazon EMR version 6.0.0.

Nouvelles fonctionnalités

- Prise en charge de l'exécution de Docker pour YARN - Les applications YARN, telles que les tâches Spark, peuvent désormais s'exécuter dans le contexte d'un conteneur Docker. Vous pouvez ainsi définir facilement des dépendances dans une image Docker sans avoir besoin d'installer des bibliothèques personnalisées sur votre cluster Amazon EMR. Pour plus d'informations, consultez Configuration de l'intégration Docker et Exécution des applications Spark avec Docker à l'aide d'Amazon EMR version 6.0.0.
- Prise en charge de LLAP pour Hive - Hive prend désormais en charge le mode d'exécution LLAP pour améliorer les performances des requêtes. Pour de plus amples informations, veuillez consulter Utilisation de LLAP pour Hive.

Modifications, améliorations et problèmes résolus

- Cette version vise à résoudre les problèmes liés à Amazon EMR Scaling lorsqu'elle ne parvient pas à up/scale réduire correctement un cluster ou entraîne des défaillances d'applications.
- Correction d'un problème où les demandes de mise à l'échelle échouaient pour un grand cluster très utilisé lorsque les démons Amazon EMR sur le cluster exécutaient des activités de surveillance de l'état, telles que la collecte de l'état des nœuds YARN et de l'état des nœuds HDFS. Cela était dû au fait que les démons du cluster n'étaient pas en mesure de communiquer les données d'état d'un nœud aux composants internes d'Amazon EMR.
- Démons EMR intégrés au cluster améliorés pour suivre correctement l'état des nœuds lorsque les adresses IP sont réutilisées afin d'améliorer la fiabilité lors des opérations de mise à l'échelle.
- SPARK-29683. Correction d'un problème où les tâches échouaient lors de la réduction de la taille du cluster, car Spark supposait que tous les nœuds disponibles étaient sur la liste de refus.

- YARN-9011. Correction d'un problème où des échecs de tâches se produisaient en raison d'une condition de course dans la mise hors service de YARN lorsque le cluster essayait d'augmenter ou de réduire sa capacité.
- Correction du problème des échecs d'étapes ou de tâches lors de la mise à l'échelle du cluster en veillant à ce que les états des nœuds soient toujours cohérents entre les démons Amazon EMR sur le cluster et YARN/HDFS.
- Correction d'un problème où les opérations de cluster telles que la réduction d'échelle et la soumission d'étapes échouaient pour les clusters Amazon EMR activés avec l'authentification Kerberos. Cela s'explique par le fait que le démon Amazon EMR intégré au cluster n'a pas renouvelé le ticket Kerberos, qui est nécessaire pour communiquer en toute sécurité avec HDFS/YARN l'exécution sur le nœud principal.
- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Amazon Linux
  - Amazon Linux 2 est le système d'exploitation de la série EMR version 6.x.
  - `systemd` est utilisé pour la gestion des services au lieu d'`upstart`, utilisé dans Amazon Linux 1.
- Kit de développement Java (JDK)
  - Corretto JDK 8 est le JDK par défaut pour la série EMR version 6.x.
- Scala
  - Scala 2.12 est utilisé avec Apache Spark et Apache Livy.
- Python 3
  - Python 3 est maintenant la version par défaut de Python dans EMR.
- Étiquettes de nœud YARN
  - À partir de la série Amazon EMR version 6.x, la fonction des étiquettes de nœud YARN est désactivée par défaut. Les processus principaux des applications peuvent s'exécuter à la fois sur les nœuds de noyau et sur les nœuds de tâche par défaut. Vous pouvez activer la fonction des étiquettes de nœud YARN en configurant les propriétés suivantes : `yarn.node-labels.enabled` et `yarn.node-labels.am.default-node-label-expression`. Pour plus d'informations, consultez Comprendre les nœuds primaires, les nœuds principaux et les nœuds de tâches.

Problèmes connus

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Redémarrer InstanceController

     ```
     $ sudo systemctl daemon-reload

     $ sudo systemctl restart instance-controller
     ```

  Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Le shell interactif Spark PySpark, y compris SparkR et spark-shell, ne prend pas en charge l'utilisation de Docker avec des bibliothèques supplémentaires.
- Pour utiliser Python 3 avec Amazon EMR version 6.0.0, vous devez ajouter `PATH` à `yarn.nodemanager.env-whitelist`.
- La fonctionnalité Live Long and Process (LLAP) n'est pas prise en charge lorsque vous utilisez le catalogue de données AWS Glue comme métastore pour Hive.
- Lorsque vous utilisez Amazon EMR 6.0.0 avec l'intégration Spark et Docker, vous devez configurer les instances de votre cluster avec le même type d'instance et la même quantité de volumes EBS pour éviter les échecs lors de la soumission d'une tâche Spark avec l'exécution Docker.
- Dans Amazon EMR 6.0.0, le mode de stockage Amazon HBase S3 est impacté par le problème HBASE-24286. HBase master ne peut pas s'initialiser lorsque le cluster est créé à l'aide de données S3 existantes.
- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

Solution :

- SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
- Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.30.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.30.1. Les modifications ont été apportées à la version 5.30.0.

Date de parution initiale : 30 juin 2020

Dernière mise à jour : 24 août 2020

Modifications, améliorations et problèmes résolus

- Les nouvelles versions d'Amazon EMR corrigent le problème en abaissant la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions d' AL2 Amazon EMR. Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent désormais un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé.
- Correction d'un problème à cause duquel le processus du contrôleur d'instance générait un nombre infini de processus.
- Correction d'un problème où Hue ne parvenait pas à exécuter une requête Hive, affichant un message « la base de données est verrouillée » et empêchant l'exécution des requêtes.
- Correction d'un problème lié à Spark qui permettait à un plus grand nombre de tâches de s'exécuter simultanément sur le cluster EMR.
- Correction d'un problème de bloc-notes Jupyter provoquant une erreur « trop de fichiers ouverts » dans le serveur Jupyter.
- Correction d'un problème lié aux temps de démarrage des clusters.

Nouvelles fonctionnalités

- Les interfaces d'application persistantes Tez UI et YARN timeline server sont disponibles avec les versions 6.x d'Amazon EMR et les versions 5.30.1 et ultérieures d'EMR. L'accès par lien en un clic à l'historique des applications persistantes vous permet d'accéder rapidement à l'historique des tâches sans avoir à configurer un proxy web par le biais d'une connexion SSH. Les journaux des clusters actifs et résiliés sont disponibles pendant 30 jours après la fin de l'application. Pour plus d'informations, consultez Affichage des interfaces utilisateur des applications persistantes dans le Guide de gestion Amazon EMR.
- L'exécution de blocs-notes EMR est disponible pour exécuter APIs des blocs-notes EMR via un script ou une ligne de commande. La possibilité de démarrer, d'arrêter, de répertorier et de décrire les exécutions d'un bloc-notes EMR sans la AWS console vous permet de contrôler un bloc-notes EMR par programme. À l'aide d'une cellule de bloc-notes paramétrée, vous pouvez transmettre différentes valeurs de paramètres à un bloc-notes sans avoir à créer une copie du bloc-notes pour chaque nouvel ensemble de valeurs de paramètres. Consultez Actions d'API EMR. Pour un exemple de code, consultez Exemples de commandes pour l'exécution de blocs-notes EMR par programmation.

Problèmes connus

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Redémarrer InstanceController

     ```
     $ sudo systemctl daemon-reload

     $ sudo systemctl restart instance-controller
     ```

  Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Blocs-notes EMR

  La fonctionnalité qui permet d'installer des noyaux et des bibliothèques Python supplémentaires sur le nœud primaire du cluster est désactivée par défaut dans la version EMR 5.30.1. Pour plus d'informations sur cette fonctionnalité, consultez Installation de noyaux et de bibliothèques Python sur un nœud primaire de cluster.

  Pour activer cette fonctionnalité, procédez comme suit :

  1. Assurez-vous que la politique d'autorisations attachée à la fonction du service pour les blocs-notes EMR autorise l'action suivante :

     ```
     elasticmapreduce:ListSteps
     ```

     Pour plus d'informations, consultez Rôle de service pour les bloc-notes EMR.

  2. Utilisez le AWS CLI pour exécuter une étape sur le cluster qui configure les Notebooks EMR, comme indiqué dans l'exemple suivant. Remplacez *us-east-1* par la région dans laquelle réside votre cluster. Pour plus d'informations sur l'ajout d'étapes, consultez la rubrique Ajout d'étapes à un cluster à l'aide de la AWS CLI.

```
aws emr add-steps  --cluster-id MyClusterID --steps
 Type=CUSTOM_JAR,Name=EMRNotebooksSetup,ActionOnFailure=CONTINUE,Jar=s3://us-
east-1.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://
awssupportdatasvcs.com/bootstrap-actions/EMRNotebooksSetup/emr-notebooks-setup.sh"]
```

- Mise à l'échelle gérée

  Les opérations de mise à l'échelle gérées sur des clusters 5.30.0 et 5.30.1 sans Presto installé peuvent provoquer des défaillances d'applications ou empêcher le maintien d'un groupe d'instances ou d'une flotte d'instances uniforme dans l'état `ARRESTED`, en particulier lorsqu'une opération de réduction est rapidement suivie d'une opération d'augementation.

  Pour contourner le problème, choisissez Presto comme application à installer lorsque vous créez un cluster avec les versions 5.30.0 et 5.30.1 d'Amazon EMR, même si votre travail ne nécessite pas Presto.

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.
  - Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère `/` (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

  Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Version 5.30.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.30.0. Les modifications ont été apportées à la version 5.29.0.

Date de parution initiale : 13 mai 2020

Date de la dernière mise à jour : 25 juin 2020

Mises à niveau

- Mise à niveau AWS SDK pour Java vers la version 1.11.759
- Mise à niveau du SDK Amazon SageMaker Spark vers la version 1.3.0
- Mise à niveau du serveur d'enregistrement EMR vers la version 1.6.0
- Mise à niveau de Flink vers la version 1.10.0
- Mise à niveau de Ganglia vers la version 3.7.2
- Mise à niveau HBase vers la version 1.4.13
- Mise à niveau de Hudi vers la version 0.5.2 incubating
- Mise à niveau de Hue vers la version 4.6.0
- Mise à niveau JupyterHub vers la version 1.1.0
- Mise à niveau de Livy vers la version 0.7.0-incubating
- Mise à niveau d'Oozie vers la version 5.2.0
- Mise à niveau de Presto vers la version 0.232
- Mise à niveau de Spark vers la version 2.4.5
- Connecteurs et pilotes mis à niveau : Amazon Glue Connector 1.12.0 ; Amazon Kinesis Connector 3.5.0 ; EMR DynamoDB Connector 4.14.0

Nouvelles fonctionnalités

- Blocs-notes EMR – Lorsqu'ils sont utilisés avec des clusters EMR créés à l'aide de la version 5.30.0, les noyaux du bloc-notes EMR s'exécutent sur un cluster. Cela améliore les performances des blocs-notes et vous permet d'installer et de personnaliser les noyaux. Vous pouvez également installer des bibliothèques Python sur le nœud primaire du cluster. Pour plus d'informations, consultez Installation et utilisation des noyaux et des bibliothèques dans le Manuel de gestion EMR .
- Dimensionnement géré – Avec les versions 5.30.0 et ultérieures d'Amazon EMR, vous pouvez activer le dimensionnement géré par EMR pour augmenter ou diminuer automatiquement le nombre d'instances ou d'unités dans votre cluster en fonction de la charge de travail. Amazon EMR évalue en permanence les métriques de cluster pour prendre des décisions de dimensionnement qui optimisent vos clusters en termes de coût et de vitesse. Pour plus d'informations, consultez Mise à l'échelle des ressources de cluster dans le Guide de gestion Amazon EMR.
- Chiffrer les fichiers journaux stockés dans Amazon S3 : avec Amazon EMR version 5.30.0 et versions ultérieures, vous pouvez chiffrer les fichiers journaux stockés dans Amazon S3 à l'aide

d'une clé gérée par le client. AWS KMS Pour plus d'informations, consultez Chiffrer les fichiers journaux stockés dans Amazon S3 dans le Guide de gestion Amazon EMR.

- Prise en charge d'Amazon Linux 2 – Les versions 5.30.0 et ultérieures d'EMR utilisent le système d'exploitation Amazon Linux 2. La nouvelle personnalisation AMIs (Amazon Machine Image) doit être basée sur l'AMI Amazon Linux 2. Pour en savoir plus, consultez Utilisation d'une image AMI personnalisée.
- Dimensionnement automatique gracieux Presto – Les clusters EMR utilisant la version 5.30.0 peuvent inclure un délai d'attente de dimensionnement automatique qui donne aux tâches Presto le temps s'exécuter complètement avant que leur nœud ne soit hors-service. Pour de plus amples informations, veuillez consulter Utilisation du dimensionnement automatique de Presto avec désaffectation gracieuse.
- Création d'une instance de flotte avec une nouvelle option de stratégie d'allocation – Une nouvelle option de stratégie d'allocation est disponible dans les versions 5.12.1 et ultérieures d'EMR. Cela permet un approvisionnement plus rapide des clusters, une allocation plus précise et moins d'interruptions d'instances Spot. Des mises à jour des rôles de service EMR autres que ceux par défaut sont requises. Consultez Configuration de parcs d'instances.
- Commandes sudo systemctl stop et sudo systemctl start – Dans les versions 5.30.0 et ultérieures d'EMR, qui utilisent le système d'exploitation Amazon Linux 2, EMR utilise les commandes `sudo systemctl stop` et `sudo systemctl start` pour redémarrer les services. Pour plus d'informations, consultez Comment redémarrer un service dans Amazon EMR ?.

Modifications, améliorations et problèmes résolus

- EMR version 5.30.0 n'installe pas Ganglia par défaut. Lorsque vous créez un cluster, vous pouvez sélectionner expressément l'installation de Ganglia.
- Optimisation des performances Spark.
- Optimisation des performances Presto.
- Python 3 est la version par défaut pour Amazon EMR version 5.30.0 et versions ultérieures.
- Le groupe de sécurité géré par défaut pour l'accès au service dans les sous-réseaux privés a été mis à jour avec de nouvelles règles. Si vous utilisez un groupe de sécurité personnalisé pour accéder au service, vous devez inclure les mêmes règles que le groupe de sécurité géré par défaut. Pour plus d'informations, consultez Groupe de sécurité géré par Amazon EMR pour l'accès au service (sous-réseaux privés). Si vous utilisez un rôle de service personnalisé pour Amazon EMR, vous devez accorder l'autorisation aux `ec2:describeSecurityGroups` pour permettre

à EMR de confirmer que les groupes de sécurité sont correctement créés. Si vous utilisez le `EMR_DefaultRole`, cette autorisation est déjà incluse dans la stratégie gérée par défaut.

Problèmes connus

- Réduction de la limite du « nombre maximum de fichiers ouverts » pour les anciennes versions AL2 [corrigée dans les nouvelles versions]. Versions Amazon EMR : emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 et emr-6.2.0 sont basées sur les anciennes versions d'Amazon Linux 2 (), qui ont un paramètre ulimit inférieur pour le « Nombre maximum de fichiers ouverts » lorsque les clusters Amazon EMR sont créés avec l'AMI par défaut. AL2 Les versions 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 et versions ultérieures d'Amazon EMR incluent un correctif permanent avec un paramètre « Nombre maximum de fichiers ouverts » plus élevé. Les versions dont la limite de fichiers ouverts est inférieure provoquent l'erreur « Trop de fichiers ouverts » lors de la soumission d'une tâche Spark. Dans les versions concernées, l'AMI par défaut Amazon EMR possède un paramètre ulimit par défaut de 4096 pour le « Nombre maximum de fichiers ouverts », ce qui est inférieur à la limite de fichiers de 65536 de la dernière AMI Amazon Linux 2. Le paramètre ulimit inférieur pour « Nombre maximum de fichiers ouverts » entraîne l'échec de la tâche Spark lorsque le pilote et l'exécuteur Spark tentent d'ouvrir plus de 4 096 fichiers. Pour résoudre ce problème, Amazon EMR dispose d'un script d'action d'amorçage (BA, bootstrap action) qui ajuste le paramètre ulimit lors de la création du cluster.

  Si vous utilisez une ancienne version d'Amazon EMR qui ne contient pas de solution permanente à ce problème, la solution suivante vous permet de définir explicitement le paramètre ulimit du contrôleur d'instance sur un maximum de 65536 fichiers.

  Définir explicitement un ulimit à partir de la ligne de commande

  1. Modifiez `/etc/systemd/system/instance-controller.service` pour ajouter les paramètres suivants à la section Service.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Redémarrer InstanceController

     ```
     $ sudo systemctl daemon-reload

     $ sudo systemctl restart instance-controller
     ```

Définissez un ulimit à l'aide de l'action d'amorçage (BA)

Vous pouvez également utiliser un script d'action d'amorçage (BA) pour configurer ulimit du contrôleur d'instance à 65536 fichiers lors de la création du cluster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Mise à l'échelle gérée

  Les opérations de mise à l'échelle gérées sur des clusters 5.30.0 et 5.30.1 sans Presto installé peuvent provoquer des défaillances d'applications ou empêcher le maintien d'un groupe d'instances ou d'une flotte d'instances uniforme dans l'état `ARRESTED`, en particulier lorsqu'une opération de réduction est rapidement suivie d'une opération d'augementation.

  Pour contourner le problème, choisissez Presto comme application à installer lorsque vous créez un cluster avec les versions 5.30.0 et 5.30.1 d'Amazon EMR, même si votre travail ne nécessite pas Presto.
- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après

que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

Solution :

- SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
- Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

- Le moteur de base de données par défaut pour Hue 4.6.0 est SQLite, ce qui pose des problèmes lorsque vous essayez d'utiliser Hue avec une base de données externe. Pour résoudre ce problème, définissez `engine` de votre configuration de classification `hue-ini` sur `mysql`. Ce problème a été résolu dans la version 5.30.1 d'Amazon EMR.
- Lorsque vous utilisez Spark avec le formatage de l'emplacement de partition Hive pour lire des données dans Amazon S3, et que vous exécutez Spark sur les versions 5.30.0 à 5.36.0 et 6.2.0 à 6.9.0 d'Amazon EMR, vous pouvez rencontrer un problème qui empêche votre cluster de lire correctement les données. Cela peut se produire si vos partitions présentent toutes les caractéristiques suivantes :
  - Deux partitions ou plus sont analysées à partir de la même table.
  - Au moins un chemin de répertoire de partition est un préfixe d'au moins un autre chemin de répertoire de partition, par exemple, `s3://bucket/table/p=a` est un préfixe de `s3://bucket/table/p=a b`.

- Le premier caractère qui suit le préfixe dans le répertoire de l'autre partition a une valeur UTF-8 inférieure au caractère / (U+002F). Par exemple, le caractère d'espace (U+0020) qui apparaît entre a et b dans `s3://bucket/table/p=a b` entre dans cette catégorie. Notez qu'il existe 14 autres caractères de non-contrôle : `!"#$%&'()*+,-`. Pour plus d'informations, consultez Table de codage UTF-8 et les caractères Unicode.

Pour contourner ce problème, définissez la configuration `spark.sql.sources.fastS3PartitionDiscovery.enabled` sur `false` dans la classification `spark-defaults`.

Version 5.29.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.29.0. Les modifications ont été apportées à la version 5.28.1.

Date de parution initiale : 17 janvier 2020

Mises à niveau

- Mise à niveau AWS SDK pour Java vers la version 1.11.682
- Mise à niveau de Hive vers la version 2.3.6
- Mise à niveau de Flink vers la version 1.9.1
- Mise à niveau d'EmrFS vers la version 2.38.0
- Mise à niveau du connecteur DynamoDB d'EMR vers la version 4.13.0

Modifications, améliorations et problèmes résolus

- Spark
  - Optimisation des performances Spark.
- EMRFS
  - Mise à jour du guide de gestion avec les paramètres par défaut du fichier emrfs-site.xml pour un affichage cohérent.

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.28.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.28.1. Les modifications ont été apportées à la version 5.28.0.

Date de parution initiale : 10 janvier 2020

Modifications, améliorations et problèmes résolus

- Spark
  - Correction des problèmes de compatibilité Spark.
- CloudWatch Métriques
  - Correction de la publication d'Amazon CloudWatch Metrics sur un cluster EMR comportant plusieurs nœuds principaux.
- Message de journal désactivé
  - Message de faux journal désactivé, « ...en utilisant l'ancienne version (<4.5.8) du client HTTP Apache ».

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.28.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.28.0. Les modifications ont été apportées à la version 5.27.0.

Date de parution initiale : 12 novembre 2019

Mises à niveau

- Mise à niveau de Flink vers la version 1.9.0
- Mise à niveau de Hive vers la version 2.3.6
- Mise à niveau MXNet vers la version 1.5.1
- Mise à niveau de Phoenix vers la version 4.14.3
- Mise à niveau de Presto vers la version 0.227
- Mise à niveau de Zeppelin vers la version 0.8.2

Nouvelles fonctionnalités

- Apache Hudi est désormais disponible pour une installation via Amazon EMR lorsque vous créez un cluster. Pour de plus amples informations, veuillez consulter Hudi.
- (25 novembre 2019) Vous pouvez maintenant choisir d'exécuter plusieurs étapes en parallèle pour améliorer l'utilisation du cluster et faire des économies. Vous pouvez également annuler à la fois les étapes en attente et celles en cours d'exécution. Pour plus d'informations, voir Travailler avec des étapes à l'aide de la console AWS CLI AND.
- (3 décembre 2019) Vous pouvez désormais créer et exécuter des clusters EMR sur. AWS Outposts AWS Outposts permet AWS des services, une infrastructure et des modèles d'exploitation natifs dans les installations sur site. Dans AWS Outposts les environnements, vous

pouvez utiliser les mêmes AWS APIs outils et infrastructures que ceux que vous utilisez dans le AWS cloud. Pour plus d'informations, consultez la section Clusters EMR sur. AWS Outposts

- (11 mars 2020) À partir de la version 5.28.0 d'Amazon EMR, vous pouvez créer et exécuter des clusters Amazon EMR sur un sous-réseau de zones AWS locales en tant qu'extension logique d'une région prenant en charge les zones locales. AWS Une zone locale permet aux fonctionnalités d'Amazon EMR et à un sous-ensemble de AWS services, tels que les services de calcul et de stockage, d'être situés plus près des utilisateurs, offrant ainsi un accès à très faible latence aux applications exécutées localement. Pour obtenir la liste des zones locales disponibles, consultez AWS Local Zones. Pour plus d'informations sur l'accès aux zones AWS locales disponibles, voir Régions, zones de disponibilité et zones locales.

  Les zones Locales ne sont actuellement pas compatibles avec Blocs-notes Amazon EMR et ne prennent pas en charge les connexions directes à Amazon EMR via le point de terminaison d'un VPC d'interface (AWS PrivateLink).

Modifications, améliorations et problèmes résolus

- Prise en charge d'applications étendues pour des clusters à haute disponibilité
  - Pour plus d'informations, consultez Applications prises en charge dans un cluster EMR avec plusieurs nœuds primaires dans le Guide de gestion Amazon EMR.
- Spark
  - Optimisation des performances
- Hive
  - Optimisation des performances
- Presto
  - Optimisation des performances

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité

du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

Solution :

- SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
- Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.27.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.27.0. Les modifications ont été apportées à la version 5.26.0.

Date de parution initiale : 23 septembre 2019

Mises à niveau

- AWS SDK pour Java 1,1,615
- Flink 1.8.1
- JupyterHub 1.0.0
- Spark 2.4.4
- TensorFlow 1.14.0
- Connecteurs et pilotes :

- Connecteur DynamoDB 4.12.0

Nouvelles fonctionnalités

- (24 octobre 2019) Les nouvelles fonctionnalités suivantes des blocs-notes EMR sont disponibles avec toutes les versions d'Amazon EMR.
  - Vous pouvez désormais associer des référentiels Git aux blocs-notes EMR pour stocker vos blocs-notes dans un environnement à version contrôlée. Vous pouvez partager du code avec vos pairs et réutiliser les blocs-notes Jupyter existants via des référentiels Git distants. Pour plus d'informations, consultez Associer des référentiels Git à Blocs-notes EMR dans le Guide de gestion Amazon EMR.
  - L'utilitaire nbdime est désormais disponible dans les blocs-notes EMR afin de simplifier la comparaison et la fusion des blocs-notes.
  - Les ordinateurs portables EMR sont désormais compatibles. JupyterLab JupyterLab est un environnement de développement interactif basé sur le Web entièrement compatible avec les ordinateurs portables Jupyter. Vous pouvez désormais choisir d'ouvrir votre bloc-notes dans l'éditeur de bloc-notes Jupyter JupyterLab ou dans l'éditeur de bloc-notes Jupyter.
- (30 octobre 2019) Avec Amazon EMR versions 5.25.0 et ultérieures, vous pouvez vous connecter à l'interface utilisateur du serveur d'historique Spark à partir de la page Récapitulatif du cluster ou de l'onglet Historique de l'application dans la console. Au lieu de configurer un proxy Web par le biais d'une connexion SSH, vous pouvez accéder rapidement à l'interface utilisateur du serveur d'historique Spark pour consulter les métriques de l'application et accéder aux fichiers journaux pertinents pour les clusters actifs et résiliés. Pour plus d'informations, consultez Accès hors cluster aux interfaces utilisateur d'application persistante dans le Guide de gestion Amazon EMR.

Modifications, améliorations et problèmes résolus

- Cluster Amazon EMR avec plusieurs nœuds primaires
  - Vous pouvez installer et exécuter Flink sur un cluster Amazon EMR comportant plusieurs nœuds primaires. Pour plus d'informations, consultez Applications et fonctionnalités prises en charge.
  - Vous pouvez configurer le chiffrement transparent HDFS sur un cluster Amazon EMR comportant plusieurs nœuds primaires. Pour plus d'informations, consultez Chiffrement transparent HDFS sur les clusters EMR comportant plusieurs nœuds primaires.

  - Vous pouvez désormais modifier la configuration des applications exécutées sur un cluster Amazon EMR comportant plusieurs nœuds primaires. Pour plus d'informations, consultez Fourniture d'une configuration pour un groupe d'instances dans un cluster en cours d'exécution.
- Connecteur Amazon EMR-DynamoDB
  - Le connecteur Amazon EMR-DynamoDB prend désormais en charge les types de données DynamoDB suivants : boolean, list, map, item, null. Pour plus d'informations, consultez Configuration d'une table Hive pour l'exécution de commandes Hive.

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos

paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.26.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.26.0. Les modifications ont été apportées à la version 5.25.0.

Date de parution initiale : 8 août 2019

Dernière mise à jour : 19 août 2019

Mises à niveau

- AWS SDK pour Java 1,1595
- HBase 1.4.10
- Phoenix 4.14.2
- Connecteurs et pilotes :
  - Connecteur DynamoDB 4.11.0
  - Connecteur MariaDB 2.4.2
  - Pilote JDBC Amazon Redshift version 1.2.32.1056

Nouvelles fonctionnalités

- (Bêta) Avec Amazon EMR 5.26.0, vous pouvez lancer un cluster qui s'intègre à Lake Formation. Cette intégration fournit un accès détaillé au niveau des colonnes aux bases de données et aux tables du Glue AWS Data Catalog. Il permet également une authentification unique fédérée à Blocs-notes EMR ou Apache Zeppelin à partir d'un système d'identité d'entreprise. Pour plus d'informations, consultez Intégrer Amazon EMR à AWS Lake Formation (version bêta).
- (19 août 2019) Bloquer l'accès public d'Amazon EMR est désormais disponible avec toutes les versions d'Amazon EMR qui prennent en charge les groupes de sécurité. Bloquer l'accès public est un paramètre applicable à l'ensemble du compte et appliqué à chaque AWS région. Le blocage de l'accès public empêche le lancement d'un cluster lorsqu'un groupe de sécurité associé au cluster possède une règle autorisant le trafic entrant depuis IPv4 0.0.0.0/0 ou IPv6 : :/0 (accès public) sur un port, sauf si un port est spécifié comme exception. Le port 22 est une exception par défaut. Pour

plus d'informations, consultez Utilisation d'Amazon EMR Block Public Access dans le Guide de gestion Amazon EMR.

Modifications, améliorations et problèmes résolus

- Blocs-notes EMR
  - Avec EMR 5.26.0 et versions ultérieures, Blocs-notes EMR prend en charge les bibliothèques Python limitées aux blocs-notes en plus des bibliothèques Python par défaut. Vous pouvez installer des bibliothèques limitées au bloc-notes depuis l'éditeur de bloc-notes sans avoir à recréer un cluster ou à rattacher un bloc-notes à un cluster. Les bibliothèques limitées aux blocs-notes sont créées dans un environnement virtuel Python, de sorte qu'elles ne s'appliquent qu'à la session de bloc-notes en cours. Cela vous permet d'isoler les dépendances du bloc-notes. Pour plus d'informations, consultez Utilisation de bibliothèques limitées aux blocs-notes dans le Guide de gestion Amazon EMR.
- EMRFS
  - Vous pouvez activer une fonctionnalité ETag de vérification (version bêta) en réglant `fs.s3.consistent.metadata.etag.verification.enabled` sur`true`. Grâce à cette fonctionnalité, EMRFS utilise Amazon S3 ETags pour vérifier que les objets lus sont de la dernière version disponible. Cette fonctionnalité est utile dans les cas read-after-update d'utilisation dans lesquels des fichiers sur Amazon S3 sont remplacés tout en conservant le même nom. Cette fonctionnalité de ETag vérification ne fonctionne pas actuellement avec S3 Select. Pour plus d'informations, consultez Configuration de la vue cohérente.
- Spark
  - Les optimisations suivantes sont désormais activées par défaut : élimination dynamique des partitions, DISTINCT avant INTERSECT, amélioration de l'inférence des statistiques de plan SQL pour les requêtes JOIN suivies de DISTINCT, aplatissement des sous-requêtes scalaires, réorganisation optimisée des jointures et jointure par filtre bloom. Pour plus d'informations, consultez Optimisation des performances de Spark.
  - Génération de code améliorée pour l'ensemble de l'étape pour Sort Merge Join.
  - Amélioration de la réutilisation des fragments de requête et des sous-requêtes.
  - Améliorations apportées à la préallocation des exécuteurs au démarrage de Spark.
  - Les jointures par filtre bloom ne sont plus appliquées lorsque le côté le plus petit de la jointure inclut un indice de diffusion.
- Tez

- Un problème avec Tez a été résolu. L'interface utilisateur de Tez fonctionne désormais sur un cluster Amazon EMR avec plusieurs nœuds primaires.

Problèmes connus

- Les fonctionnalités améliorées de génération de code pour l'ensemble de l'étape pour Sort Merge Join peuvent augmenter la sollicitation de la mémoire lorsqu'elles sont activées. Cette optimisation améliore les performances, mais peut entraîner de nouvelles tentatives de tâches ou des échecs si `spark.yarn.executor.memoryOverheadFactor` n'est pas réglé pour fournir suffisamment de mémoire. Pour désactiver cette fonctionnalité, définissez `spark.sql.sortMergeJoinExec.extendedCodegen.enabled` sur false.
- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos

paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.25.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.25.0. Les modifications ont été apportées à la version 5.24.1.

Date de parution initiale : 17 juillet 2019

Dernière mise à jour : 30 octobre 2019

Amazon EMR 5.25.0

Mises à niveau

- AWS SDK pour Java 1,1,566
- Hive 2.3.5
- Presto 0.220
- Spark 2.4.3
- TensorFlow 1.13.1
- Tez 0.9.2
- Zookeeper 3.4.14

Nouvelles fonctionnalités

- (30 octobre 2019) À partir de la version 5.25.0 d'Amazon EMR, vous pouvez vous connecter à l'interface utilisateur du serveur d'historique Spark à partir de la page Récapitulatif du cluster ou de l'onglet Historique de l'application dans la console. Au lieu de configurer un proxy Web par le biais d'une connexion SSH, vous pouvez accéder rapidement à l'interface utilisateur du serveur d'historique Spark pour consulter les métriques de l'application et accéder aux fichiers journaux pertinents pour les clusters actifs et résiliés. Pour plus d'informations, consultez Accès hors cluster aux interfaces utilisateur d'application persistante dans le Guide de gestion Amazon EMR.

Modifications, améliorations et problèmes résolus

- Spark

- Amélioration des performances de certaines jointures en utilisant des filtres Bloom pour préfiltrer les entrées. L'optimisation est désactivée par défaut et peut être activée en définissant le paramètre de configuration Spark `spark.sql.bloomFilterJoin.enabled` sur `true`.
- Amélioration des performances du regroupement par colonnes de type chaîne.
- Amélioration de la mémoire par défaut de l'exécuteur Spark et de la configuration des cœurs des types d'instances R4 pour les clusters non HBase installés.
- Résolution d'un problème antérieur lié à la fonctionnalité d'élimination dynamique des partitions, où la table éliminée devait se trouver du côté gauche de la jointure.
- Amélioration de l'optimisation DISTINCT avant INTERSECT pour l'appliquer à des cas supplémentaires impliquant des alias.
- Amélioration de l'inférence des statistiques de plan SQL pour les requêtes JOIN suivies de DISTINCT. Cette amélioration est désactivée par défaut et peut être activée en définissant le paramètre de configuration Spark `spark.sql.statsImprovements.enabled` sur `true`. Cette optimisation est requise par la fonctionnalité Distinct avant Intersect et sera automatiquement activée lorsque `spark.sql.optimizer.distinctBeforeIntersect.enabled` est définie sur `true`.
- Ordre de jointure optimisé en fonction de la taille de la table et des filtres. Cette optimisation est désactivée par défaut et peut être activée en définissant le paramètre de configuration Spark `spark.sql.optimizer.sizeBasedJoinReorder.enabled` sur `true`.

  Pour plus d'informations, consultez Optimisation des performances de Spark.

- EMRFS
  - Le paramètre EMRFS, `fs.s3.buckets.create.enabled`, est désormais désactivé par défaut. Lors des tests, nous avons constaté que la désactivation de ce paramètre améliore les performances et empêche la création involontaire de compartiments S3. Si votre application repose sur cette fonctionnalité, vous pouvez l'activer en définissant la propriété `fs.s3.buckets.create.enabled` sur `true` dans la classification de configuration `emrfs-site`. Pour plus d'informations, consultez Fourniture d'une configuration lors de la création d'un cluster.
- Améliorations du chiffrement du disque local et du chiffrement S3 dans les configurations de sécurité (5 août 2019)
  - Les paramètres de chiffrement Amazon S3 ont été séparés des paramètres de chiffrement du disque local dans la configuration de la sécurité.

- Ajout d'une option permettant d'activer le chiffrement EBS avec les versions 5.24.0 et ultérieures. La sélection de cette option chiffre le volume du périphérique racine en plus des volumes de stockage. Les versions précédentes nécessitaient l'utilisation d'une AMI personnalisée pour chiffrer le volume du périphérique racine.
- Pour plus d'informations, consultez Options de chiffrement dans le Guide de gestion Amazon EMR.

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.24.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.24.1. Les modifications ont été apportées à la version 5.24.0.

Date de parution initiale : 26 juin 2019

Modifications, améliorations et problèmes résolus

- Mise à jour de l'AMI Amazon Linux par défaut pour Amazon EMR afin d'inclure d'importantes mises à jour de sécurité du noyau Linux, notamment le problème de déni de service TCP SACK (AWS-2019-005).

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.24.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.24.0. Les modifications ont été apportées à la version 5.23.0.

Date de parution initiale : 11 juin 2019

Dernière mise à jour : 5 août 2019

Mises à niveau

- Flink 1.8.0
- Hue 4.4.0
- JupyterHub 0,9,6
- Livy 0.6.0
- MxNet 1.4.0
- Presto 0.219
- Spark 2.4.2
- AWS SDK pour Java 1,1,546
- Connecteurs et pilotes :
  - Connecteur DynamoDB 4.9.0
  - Connecteur MariaDB 2.4.1
  - Pilote JDBC Amazon Redshift version 1.2.27.1051

Modifications, améliorations et problèmes résolus

- Spark

- Ajout d'une optimisation pour éliminer dynamiquement les partitions. L'optimisation est désactivée par défaut. Pour l'activer, définissez le paramètre de configuration Spark `spark.sql.dynamicPartitionPruning.enabled` sur `true`.
- Amélioration des performances des requêtes `INTERSECT`. L'optimisation est désactivée par défaut. Pour l'activer, définissez le paramètre de configuration Spark `spark.sql.optimizer.distinctBeforeIntersect.enabled` sur `true`.
- Ajout d'une optimisation pour aplatir les sous-requêtes scalaires avec des agrégats qui utilisent la même relation. L'optimisation est désactivée par défaut. Pour l'activer, définissez le paramètre de configuration Spark `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled` sur `true`.
- Génération de code améliorée pour l'ensemble de l'étape.

Pour plus d'informations, consultez Optimisation des performances de Spark.

- Améliorations du chiffrement du disque local et du chiffrement S3 dans les configurations de sécurité (5 août 2019)
  - Les paramètres de chiffrement Amazon S3 ont été séparés des paramètres de chiffrement du disque local dans la configuration de la sécurité.
  - Ajout d'une option pour activer le chiffrement EBS. La sélection de cette option chiffre le volume du périphérique racine en plus des volumes de stockage. Les versions précédentes nécessitaient l'utilisation d'une AMI personnalisée pour chiffrer le volume du périphérique racine.
  - Pour plus d'informations, consultez Options de chiffrement dans le Guide de gestion Amazon EMR.

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois

sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

Solution :

- SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
- Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.23.0. Les modifications ont été apportées à la version 5.22.0.

Date de parution initiale : 1 avril 2019

Dernière mise à jour : 30 avril 2019

Mises à niveau

- AWS SDK pour Java 1,1,519

Nouvelles fonctionnalités

- (30 avril 2019) Avec Amazon EMR 5.23.0 et versions ultérieures, vous pouvez lancer un cluster avec trois nœuds principaux pour prendre en charge la haute disponibilité d'applications telles que YARN Resource Manager, HDFS, Spark NameNode, Hive et Ganglia. Le nœud primaire n'est plus

un point de défaillance potentiel grâce à cette fonctionnalité. Si l'un des nœuds primaires tombe en panne, Amazon EMR passe automatiquement sur un nœud primaire de secours et remplace le nœud primaire défaillant par un nouveau nœud ayant la même configuration et les mêmes actions de démarrage. Pour plus d'informations, consultez Planification et configuration des nœuds primaires.

Problèmes connus

- Interface utilisateur Tez (corrigée dans la version 5.26.0 d'Amazon EMR)

  L'interface utilisateur de Tez ne fonctionne pas sur un cluster EMR comportant plusieurs nœuds primaires.

- Hue (corrigée dans la version 5.24.0 d'Amazon EMR)
  - Hue exécuté sur Amazon EMR ne prend pas en charge Solr. À compter de la version 5.20.0 d'Amazon EMR, un problème de configuration incorrect entraîne l'activation de Solr et un message d'erreur inoffensif semblable au suivant s'affiche :

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Pour empêcher l'affichage du message d'erreur de Solr, procédez comme suit :

    1. Connectez-vous à la ligne de commande du nœud primaire à l'aide de SSH.
    2. Utilisez un éditeur de texte pour ouvrir le fichier `hue.ini`. Par exemple :

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```

    3. Recherchez le terme `appblacklist` et modifiez la ligne comme suit :

       ```
       appblacklist = search
       ```

    4. Enregistrez vos modifications et redémarrez Hue comme indiqué dans l'exemple suivant :

       ```
       sudo stop hue; sudo start hue
       ```

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.22.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.22.0. Les modifications ont été apportées à la version 5.21.0.

> **Important**
>
> Depuis la version 5.22.0 d'Amazon EMR, Amazon EMR AWS utilise Signature version 4 exclusivement pour authentifier les demandes adressées à Amazon S3. Les versions

antérieures d'Amazon EMR utilisent la version 2 de AWS Signature dans certains cas, sauf si les notes de publication indiquent que la version 4 de Signature est utilisée exclusivement. Pour plus d'informations, consultez les sections Authentification des demandes (AWS Signature version 4) et Authentification des demandes (AWS Signature version 2) dans le manuel Amazon Simple Storage Service Developer Guide.

Date de parution initiale : 20 mars 2019

Mises à niveau

- Flink 1.7.1
- HBase 1.4.9
- Oozie 5.1.0
- Phoenix 4.14.1
- Zeppelin 0.8.1
- Connecteurs et pilotes :
  - Connecteur DynamoDB 4.8.0
  - Connecteur MariaDB 2.2.6
  - Pilote JDBC Amazon Redshift version 1.2.20.1043

Nouvelles fonctionnalités

- Modification de la configuration EBS par défaut pour les types d' EC2 instances avec stockage EBS uniquement. Lorsque vous créez un cluster à l'aide d'Amazon EMR version 5.22.0 et versions ultérieures, la quantité de stockage EBS par défaut augmente en fonction de la taille de l'instance. En outre, nous avons fractionné une plus grande quantité de stockage sur plusieurs volumes, afin d'offrir de meilleures performances en termes d'IOPS. Si vous voulez utiliser une autre configuration de stockage d'instance EBS, vous pouvez la spécifier lorsque vous créez un cluster EMR ou ajoutez des nœuds à un cluster existant. Pour plus d'informations sur la quantité de stockage et le nombre de volumes alloués par défaut pour chaque type d'instance, consultez Stockage EBS par défaut pour les instances dans le Guide de gestion Amazon EMR.

Modifications, améliorations et problèmes résolus

- Spark

  - Introduction d'une nouvelle propriété de configuration pour Spark sur YARN, `spark.yarn.executor.memoryOverheadFactor`. La valeur de cette propriété est un facteur d'échelle qui définit la valeur de la surcharge mémoire à un pourcentage de la mémoire de l'exécuteur, avec un minimum de 384 Mo. Si la surcharge mémoire est définie explicitement à l'aide `spark.yarn.executor.memoryOverhead`, cette propriété n'a aucun effet. La valeur par défaut est `0.1875`, ce qui représente 18,75 %. Cette valeur par défaut pour Amazon EMR laisse plus d'espace dans les conteneurs YARN pour la surcharge de mémoire de l'exécuteur que la valeur par défaut de 10 % définie en interne par Spark. La valeur par défaut d'Amazon EMR de 18,75 % a montré de façon empirique moins de défaillances liées à la mémoire dans les évaluations comparatives TPC-DS.
  - Rétroportage de SPARK-26316 pour améliorer les performances.
- Dans les versions 5.19.0, 5.20.0 et 5.21.0 d'Amazon EMR, les étiquettes des nœuds YARN sont stockées dans un répertoire HDFS. Dans certaines situations, cela entraîne des retards dans le démarrage des nœuds principaux, ce qui provoque le dépassement du délai du cluster et l'échec du lancement. À partir d'Amazon EMR 5.22.0, ce problème est résolu. Les étiquettes des nœuds YARN sont stockées sur le disque local de chaque nœud du cluster, évitant ainsi toute dépendance vis-à-vis du HDFS.

Problèmes connus

- Hue (corrigée dans la version 5.24.0 d'Amazon EMR)
  - Hue exécuté sur Amazon EMR ne prend pas en charge Solr. À compter de la version 5.20.0 d'Amazon EMR, un problème de configuration incorrect entraîne l'activation de Solr et un message d'erreur inoffensif semblable au suivant s'affiche :

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Pour empêcher l'affichage du message d'erreur de Solr, procédez comme suit :

    1. Connectez-vous à la ligne de commande du nœud primaire à l'aide de SSH.
    2. Utilisez un éditeur de texte pour ouvrir le fichier `hue.ini`. Par exemple :

```
sudo vim /etc/hue/conf/hue.ini
```

3. Recherchez le terme `appblacklist` et modifiez la ligne comme suit :

```
appblacklist = search
```

4. Enregistrez vos modifications et redémarrez Hue comme indiqué dans l'exemple suivant :

```
sudo stop hue; sudo start hue
```

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

  Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.21.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.21.1. Les modifications ont été apportées à la version 5.21.0.

Date de parution initiale : 18 juillet 2019

Modifications, améliorations et problèmes résolus

- Mise à jour de l'AMI Amazon Linux par défaut pour Amazon EMR afin d'inclure d'importantes mises à jour de sécurité du noyau Linux, notamment le problème de déni de service TCP SACK (AWS-2019-005).

Problèmes connus

- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.21.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.21.0. Les modifications ont été apportées à la version 5.20.0.

Date de parution initiale : 18 février 2019

Dernière mise à jour : 3 avril 2019

Mises à niveau

- Flink 1.7.0
- Presto 0.215
- AWS SDK pour Java 1,11,479

Nouvelles fonctionnalités

- (3 avril 2019) Avec la version 5.21.0 et ultérieures d'Amazon EMR, vous permet de remplacer les configurations de cluster et de spécifier des classifications de configuration supplémentaires pour chaque groupe d'instances dans un cluster en cours d'exécution. Pour ce faire, utilisez la console Amazon EMR, le AWS Command Line Interface (AWS CLI) ou le AWS SDK. Pour plus d'informations, consultez Fourniture d'une configuration pour un groupe d'instances dans un cluster en cours d'exécution.

Modifications, améliorations et problèmes résolus

- Zeppelin
  - Rétroportage de ZEPPELIN-3878.

Problèmes connus

- Hue (corrigée dans la version 5.24.0 d'Amazon EMR)
  - Hue exécuté sur Amazon EMR ne prend pas en charge Solr. À compter de la version 5.20.0 d'Amazon EMR, un problème de configuration incorrect entraîne l'activation de Solr et un message d'erreur inoffensif semblable au suivant s'affiche :

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Pour empêcher l'affichage du message d'erreur de Solr, procédez comme suit :

    1. Connectez-vous à la ligne de commande du nœud primaire à l'aide de SSH.
    2. Utilisez un éditeur de texte pour ouvrir le fichier `hue.ini`. Par exemple :

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```
    3. Recherchez le terme `appblacklist` et modifiez la ligne comme suit :

       ```
       appblacklist = search
       ```
    4. Enregistrez vos modifications et redémarrez Hue comme indiqué dans l'exemple suivant :

       ```
       sudo stop hue; sudo start hue
       ```
- Tez
  - Ce problème a été corrigé dans Amazon EMR 5.22.0.

    Lorsque vous vous connectez à l'interface utilisateur de Tez à l'adresse http : //:8080/tez-ui *`MasterDNS`* via une connexion SSH au nœud principal du cluster, l'erreur « L'opération de l'adaptateur a échoué - Le serveur Timeline (ATS) est hors de portée. Soit il est en panne, soit CORS n'est pas activé » apparaît, ou les tâches affichent de manière inattendue « N/A ».

    Cela est dû au fait que l'interface utilisateur de Tez envoie des demandes à YARN Timeline Server en utilisant `localhost` plutôt que le nom d'hôte du nœud primaire. Pour contourner le problème, un script est disponible pour être exécuté en tant qu'action ou étape d'amorçage. Le

script met à jour le nom d'hôte dans le fichier Tez `configs.env`. Pour plus d'informations et pour connaître l'emplacement du script, consultez les Instructions sur l'amorçage.

- Dans les versions 5.19.0, 5.20.0 et 5.21.0 d'Amazon EMR, les étiquettes des nœuds YARN sont stockées dans un répertoire HDFS. Dans certaines situations, cela entraîne des retards dans le démarrage des nœuds principaux, ce qui provoque le dépassement du délai du cluster et l'échec du lancement. À partir d'Amazon EMR 5.22.0, ce problème est résolu. Les étiquettes des nœuds YARN sont stockées sur le disque local de chaque nœud du cluster, évitant ainsi toute dépendance vis-à-vis du HDFS.
- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

  Solution :

  - SSH en tant qu'utilisateur `hadoop` au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
  - Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Généralement, le fichier keytab se trouve dans `/etc/hadoop.keytab` et le principal se présente sous la forme de `hadoop/<hostname>@<REALM>`.

    > **Note**
    >
    > Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.20.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.20.0. Les modifications ont été apportées à la version 5.19.0.

Date de parution initiale : 18 décembre 2018

Dernière mise à jour : 22 janvier 2019

Mises à niveau

- Flink 1.6.2
- HBase 1.4.8
- Hive 2.3.4
- Hue 4.3.0
- MXNet 1.3.1
- Presto 0.214
- Spark 2.4.0
- TensorFlow 1,12,0
- Tez 0.9.1
- AWS SDK pour Java 1,11,461

Nouvelles fonctionnalités

- (22 janvier 2019) Kerberos dans Amazon EMR a été amélioré pour prendre en charge l'authentification des principaux à partir d'un KDC externe. Cela permet de centraliser la gestion des principaux, car plusieurs clusters peuvent partager un seul KDC externe. De plus, le KDC externe peut avoir une relation d'approbation inter-domaines avec un domaine Active Directory. Cela permet à tous les clusters d'authentifier les mandataires à partir d'Active Directory. Pour plus d'informations, consultez Utilisation de l'authentification Kerberos dans le Guide de gestion Amazon EMR.

Modifications, améliorations et problèmes résolus

- AMI Amazon Linux par défaut pour Amazon EMR
  - Le package Python3 a été mis à jour de python 3.4 à 3.6.

- Le valideur EMRFS optimisé pour S3
  - Le validateur EMRFS optimisé pour S3 est désormais activé par défaut, ce qui améliore les performances d'écriture. Pour de plus amples informations, veuillez consulter Utilisation d'un valideur EMRFS optimisé pour S3.
- Hive
  - Rétroportage de HIVE-16686.
- Glue avec Spark et Hive
  - Dans EMR 5.20.0 ou version ultérieure, l'élagage des partitions parallèles est activé automatiquement pour Spark et Hive lorsque AWS Glue Data Catalog est utilisé comme métastore. Cette modification réduit considérablement le temps de planification des requêtes en exécutant plusieurs requêtes en parallèle pour récupérer des partitions. Le nombre total de segments pouvant être exécutés simultanément est compris entre 1 et 10. La valeur par défaut est 5, ce qui est recommandé. Vous pouvez le modifier en spécifiant la propriété `aws.glue.partition.num.segments` dans la classification de configuration `hive-site`. En cas de limitation, vous pouvez désactiver la fonctionnalité en remplaçant la valeur par 1. Pour en savoir plus, consultez Structure d'un segment AWS Glue.

Problèmes connus

- Hue (corrigée dans la version 5.24.0 d'Amazon EMR)
  - Hue exécuté sur Amazon EMR ne prend pas en charge Solr. À compter de la version 5.20.0 d'Amazon EMR, un problème de configuration incorrect entraîne l'activation de Solr et un message d'erreur inoffensif semblable au suivant s'affiche :

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Pour empêcher l'affichage du message d'erreur de Solr, procédez comme suit :

    1. Connectez-vous à la ligne de commande du nœud primaire à l'aide de SSH.
    2. Utilisez un éditeur de texte pour ouvrir le fichier `hue.ini`. Par exemple :

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```

3. Recherchez le terme `appblacklist` et modifiez la ligne comme suit :

```
appblacklist = search
```

4. Enregistrez vos modifications et redémarrez Hue comme indiqué dans l'exemple suivant :

```
sudo stop hue; sudo start hue
```

- Tez
  - Ce problème a été corrigé dans Amazon EMR 5.22.0.

    Lorsque vous vous connectez à l'interface utilisateur de Tez à l'adresse http : //:8080/tez-ui *MasterDNS* via une connexion SSH au nœud principal du cluster, l'erreur « L'opération de l'adaptateur a échoué - Le serveur Timeline (ATS) est hors de portée. Soit il est en panne, soit CORS n'est pas activé » apparaît, ou les tâches affichent de manière inattendue « N/A ».

    Cela est dû au fait que l'interface utilisateur de Tez envoie des demandes à YARN Timeline Server en utilisant `localhost` plutôt que le nom d'hôte du nœud primaire. Pour contourner le problème, un script est disponible pour être exécuté en tant qu'action ou étape d'amorçage. Le script met à jour le nom d'hôte dans le fichier Tez `configs.env`. Pour plus d'informations et pour connaître l'emplacement du script, consultez les Instructions sur l'amorçage.
- Dans les versions 5.19.0, 5.20.0 et 5.21.0 d'Amazon EMR, les étiquettes des nœuds YARN sont stockées dans un répertoire HDFS. Dans certaines situations, cela entraîne des retards dans le démarrage des nœuds principaux, ce qui provoque le dépassement du délai du cluster et l'échec du lancement. À partir d'Amazon EMR 5.22.0, ce problème est résolu. Les étiquettes des nœuds YARN sont stockées sur le disque local de chaque nœud du cluster, évitant ainsi toute dépendance vis-à-vis du HDFS.
- Problème connu dans les clusters dotés de plusieurs nœuds primaires et d'une authentification Kerberos

  Si vous exécutez des clusters avec plusieurs nœuds primaires et une authentification Kerberos dans les versions 5.20.0 et ultérieures d'Amazon EMR, vous pouvez rencontrer des problèmes avec des opérations de cluster telles que la réduction d'échelle ou la soumission d'étapes, après que le cluster ait fonctionné pendant un certain temps. La durée dépend de la période de validité du ticket Kerberos que vous avez définie. Le problème de réduction d'échelle a un impact à la fois sur la réduction d'échelle automatique et sur les demandes de réduction d'échelle explicites que vous avez soumises. D'autres opérations de cluster peuvent également être affectées.

Solution :

- SSH en tant qu'utilisateur hadoop au nœud primaire du cluster EMR avec plusieurs nœuds primaires.
- Exécutez la commande suivante pour renouveler le ticket Kerberos pour l'utilisateur hadoop.

```
kinit -kt <keytab_file> <principal>
```

Généralement, le fichier keytab se trouve dans /etc/hadoop.keytab et le principal se présente sous la forme de hadoop/<hostname>@<REALM>.

> **Note**
>
> Cette solution de contournement sera effective pendant toute la durée de validité du ticket Kerberos. Cette durée est de 10 heures par défaut, mais peut être configurée par vos paramètres Kerberos. Vous devez exécuter à nouveau la commande ci-dessus une fois le ticket Kerberos expiré.

Version 5.19.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.19.0. Les modifications ont été apportées à la version 5.18.0.

Date de parution initiale : 7 novembre 2018

Date de la dernière mise à jour : 19 novembre 2018

Mises à niveau

- Hadoop 2.8.5
- Flink 1.6.1
- JupyterHub 0,9.4
- MXNet 1.3.0
- Presto 0.212
- TensorFlow 1,11,0
- Zookeeper 3.4.13

- AWS SDK pour Java 1,1,433

Nouvelles fonctionnalités

- (19 novembre 2018) Blocs-notes EMR est un environnement géré basé sur le bloc-notes Jupyter. Il prend en charge les noyaux magiques Spark pour PySpark Spark SQL, Spark R et Scala. Blocs-notes EMR peut être utilisé avec des clusters créés à l'aide de la version 5.18.0 d'Amazon EMR ou d'une version ultérieure. Pour plus d'informations, consultez Utilisation de Blocs-notes EMR dans le Guide de gestion d'Amazon EMR.
- Le validateur EMRFS optimisé pour S3 est disponible lors de l'écriture de fichiers Parquet avec Spark et EMRFS. Ce validateur améliore les performances d'écriture. Pour de plus amples informations, veuillez consulter Utilisation d'un valideur EMRFS optimisé pour S3.

Modifications, améliorations et problèmes résolus

- YARN
  - Modification de la logique qui limite le processus principal de l'application à l'exécution sur les nœuds principaux. Cette fonctionnalité utilise désormais la fonctionnalité et les propriétés des étiquettes de nœuds YARN dans les classifications de configuration `yarn-site` et `capacity-scheduler`. Pour plus d'informations, veuillez consulter https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-plan-instances-guidelines.html#emr-plan-spot-YARN.
- AMI Amazon Linux par défaut pour Amazon EMR
  - `ruby18`, `php56` et `gcc48` ne sont plus installés par défaut. Ils peuvent être installés si vous le souhaitez à l'aide de `yum`.
  - Le gem ruby aws-sdk n'est plus installé par défaut. Il peut être installé en utilisant `gem install aws-sdk`, si vous le souhaitez. Des composants spécifiques peuvent également être installés. Par exemple, `gem install aws-sdk-s3`.

Problèmes connus

- Blocs-notes EMR – Dans certains cas, lorsque plusieurs éditeurs de blocs-notes sont ouverts, l'éditeur de bloc-notes peut sembler incapable de se connecter au cluster. Dans ce cas, effacez les cookies du navigateur, puis rouvrez les éditeurs de bloc-notes.
- CloudWatch ContainerPending Mise à l'échelle métrique et automatique — (Corrigé dans la version 5.20.0) Amazon EMR peut émettre une valeur négative pour. `ContainerPending`

Si `ContainerPending` est utilisée dans une règle de mise à l'échelle automatique, la mise à l'échelle automatique ne se comporte pas comme prévu. Évitez d'utiliser `ContainerPending` avec la mise à l'échelle automatique.

- Dans les versions 5.19.0, 5.20.0 et 5.21.0 d'Amazon EMR, les étiquettes des nœuds YARN sont stockées dans un répertoire HDFS. Dans certaines situations, cela entraîne des retards dans le démarrage des nœuds principaux, ce qui provoque le dépassement du délai du cluster et l'échec du lancement. À partir d'Amazon EMR 5.22.0, ce problème est résolu. Les étiquettes des nœuds YARN sont stockées sur le disque local de chaque nœud du cluster, évitant ainsi toute dépendance vis-à-vis du HDFS.

Version 5.18.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.18.0. Les modifications ont été apportées à la version 5.17.0.

Date de parution initiale : 24 octobre 2018

Mises à niveau

- Flink 1.6.0
- HBase 1.4.7
- Presto 0.210
- Spark 2.3.2
- Zeppelin 0.8.0

Nouvelles fonctionnalités

- À partir d'Amazon EMR 5.18.0, vous pouvez utiliser le référentiel d'artefacts d'Amazon EMR pour générer le code de votre tâche en fonction des versions exactes des bibliothèques et des dépendances qui sont disponibles avec des versions spécifiques d'Amazon EMR. Pour de plus amples informations, veuillez consulter Vérification des dépendances à l'aide du référentiel d'artefacts d'Amazon EMR.

Modifications, améliorations et problèmes résolus

- Hive

  - Ajout de la prise en charge de S3 Select. Pour de plus amples informations, veuillez consulter Utilisation de S3 Select avec Hive pour améliorer les performances.
- Presto
  - Ajout de la prise en charge de S3 Select Pushdown. Pour de plus amples informations, veuillez consulter Utilisation de S3 Select Pushdown avec Presto pour améliorer les performances.
- Spark
  - La configuration par défaut de log4j pour Spark a été modifiée pour enregistrer les journaux des conteneurs toutes les heures pour les tâches de streaming Spark. Cela permet d'éviter la suppression des journaux pour les tâches de streaming Spark de longue durée.

Version 5.17.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.17.1. Les modifications ont été apportées à la version 5.17.0.

Date de parution initiale : 18 juillet 2019

Modifications, améliorations et problèmes résolus

- Mise à jour de l'AMI Amazon Linux par défaut pour Amazon EMR afin d'inclure d'importantes mises à jour de sécurité du noyau Linux, notamment le problème de déni de service TCP SACK (AWS-2019-005).

Version 5.17.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.17.0. Les modifications ont été apportées à la version 5.16.0.

Date de parution initiale : 30 août 2018

Mises à niveau

- Flink 1.5.2
- HBase 1.4.6
- Presto 0.206

Nouvelles fonctionnalités

- Ajout de la prise en charge de Tensorflow. Pour de plus amples informations, veuillez consulter TensorFlow.

Modifications, améliorations et problèmes résolus

- JupyterHub
  - Ajout de la prise en charge de la persistance des blocs-notes dans Amazon S3. Pour de plus amples informations, veuillez consulter Configuration de la persistance pour les blocs-notes dans Amazon S3.
- Spark
  - Ajout de la prise en charge de S3 Select. Pour de plus amples informations, veuillez consulter Utilisation de S3 Select avec Spark pour améliorer les performances des requêtes.
- Résolution des problèmes liés aux métriques Cloudwatch et à la fonctionnalité de mise à l'échelle automatique dans Amazon EMR version 5.14.0, 5.15.0 ou 5.16.0.

Problèmes connus

- Lorsque vous créez un cluster activé pour Kerberos avec Livy installé, Livy échoue avec un message d'erreur indiquant que l'authentification simple n'est pas activée. Le redémarrage du serveur Livy résout le problème. Pour contourner le problème, ajoutez une étape lors de la création du cluster qui exécute `sudo restart livy-server` sur le nœud primaire.
- Si vous utilisez une AMI Amazon Linux personnalisée basée sur une AMI Amazon Linux dont la date de création est le 11/08/2018, le serveur Oozie ne démarre pas. Si vous utilisez Oozie, créez une AMI personnalisée basée sur un ID d'AMI Amazon Linux avec une date de création différente. Vous pouvez utiliser la AWS CLI commande suivante pour renvoyer une liste d'images IDs pour tous les HVM Amazon Linux AMIs avec une version 2018.03, ainsi que la date de sortie, afin de pouvoir choisir une AMI Amazon Linux appropriée comme base. MyRegion Remplacez-le par votre identifiant de région, tel que us-west-2.

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`]|[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

Version 5.16.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.16.0. Les modifications ont été apportées à la version 5.15.0.

Date de parution initiale : 19 juillet 2018

Mises à niveau

- Hadoop 2.8.4
- Flink 1.5.0
- Livy 0.5.0
- MXNet 1.2.0
- Phoenix 4.14.0
- Presto 0.203
- Spark 2.3.1
- AWS SDK pour Java 1,1336
- CUDA 9.2
- Pilote JDBC Redshift 1.2.15.1025

Modifications, améliorations et problèmes résolus

- HBase
  - Rétroportage de HBASE-20723.
- Presto
  - Modifications de configuration pour prendre en charge l'authentification LDAP. Pour de plus amples informations, veuillez consulter Utilisation de l'authentification LDAP pour Presto sur Amazon EMR.
- Spark
  - Apache Spark version 2.3.1, disponible à partir de la version 5.16.0 d'Amazon EMR, corrige CVE-2018-8024 et CVE-2018-1334. Nous vous recommandons de migrer les versions antérieures de Spark vers la version 2.3.1 ou ultérieure.

Problèmes connus

- Cette version ne prend pas en charge les types d'instance c1.medium ou m1.small. Les clusters utilisant l'un ou l'autre de ces types d'instances ne démarrent pas. Pour contourner le problème, spécifiez un type d'instance différent ou utilisez une version différente.
- Lorsque vous créez un cluster activé pour Kerberos avec Livy installé, Livy échoue avec un message d'erreur indiquant que l'authentification simple n'est pas activée. Le redémarrage du serveur Livy résout le problème. Pour contourner le problème, ajoutez une étape lors de la création du cluster qui exécute `sudo restart livy-server` sur le nœud primaire.
- Après le redémarrage du nœud principal ou le redémarrage du contrôleur d'instance, les CloudWatch métriques ne seront pas collectées et la fonctionnalité de dimensionnement automatique ne sera pas disponible dans les versions 5.14.0, 5.15.0 ou 5.16.0 d'Amazon EMR. Ce problème est résolu dans Amazon EMR 5.17.0.

Version 5.15.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.15.0. Les modifications ont été apportées à la version 5.14.0.

Date de parution initiale : 21 juin 2018

Mises à niveau

- Mise à niveau HBase vers la version 1.4.4
- Mise à niveau de Hive vers la version 2.3.3
- Mise à niveau de Hue vers la version 4.2.0
- Mise à niveau d'Oozie vers la version 5.0.0
- Mise à niveau de Zookeeper vers la version 3.4.12
- AWS SDK mis à jour vers la version 1.11.333

Modifications, améliorations et problèmes résolus

- Hive
  - Rétroportage de HIVE-18069.
- Hue

  - Hue a été mis à jour pour s'authentifier correctement auprès de Livy lorsque Kerberos est activé. Livy est désormais pris en charge lors de l'utilisation de Kerberos avec Amazon EMR.
- JupyterHub
  - Mis à jour JupyterHub afin qu'Amazon EMR installe les bibliothèques clientes LDAP par défaut.
  - Correction d'une erreur dans le script qui génère des certificats auto-signés.

Problèmes connus

- Cette version ne prend pas en charge les types d'instance c1.medium ou m1.small. Les clusters utilisant l'un ou l'autre de ces types d'instances ne démarrent pas. Pour contourner le problème, spécifiez un type d'instance différent ou utilisez une version différente.
- Après le redémarrage du nœud principal ou le redémarrage du contrôleur d'instance, les CloudWatch métriques ne seront pas collectées et la fonctionnalité de dimensionnement automatique ne sera pas disponible dans les versions 5.14.0, 5.15.0 ou 5.16.0 d'Amazon EMR. Ce problème est résolu dans Amazon EMR 5.17.0.

Version 5.14.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.14.1. Les modifications ont été apportées à la version 5.14.0.

Date de parution initiale : 17 octobre 2018

Mise à jour de l'AMI par défaut pour Amazon EMR afin de corriger les vulnérabilités de sécurité potentielles.

Version 5.14.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.14.0. Les modifications ont été apportées à la version 5.13.0.

Date de parution initiale : 4 juin 2018

Mises à niveau

- Mise à niveau d'Apache Flink vers la version 1.4.2
- Mise à niveau d'Apache MXnet vers la version 1.1.0

- Mise à niveau d'Apache Sqoop vers la version 1.4.7

Nouvelles fonctionnalités

- JupyterHub Support ajouté. Pour de plus amples informations, veuillez consulter JupyterHub.

Modifications, améliorations et problèmes résolus

- EMRFS
  - La chaîne userAgent dans les demandes adressées à Amazon S3 a été mise à jour pour contenir les informations relatives à l'utilisateur et au groupe du principal invoquant. Cela peut être utilisé avec AWS CloudTrail les journaux pour un suivi plus complet des demandes.
- HBase
  - Inclusion de HBASE-20447, qui corrige un problème qui pouvait causer des problèmes de cache, en particulier avec les régions divisées.
- MXnet
  - Ajout des bibliothèques OpenCV.
- Spark
  - Lorsque Spark écrit des fichiers Parquet sur un emplacement Amazon S3 à l'aide d'EMRFS, l' FileOutputCommitter algorithme a été mis à jour pour utiliser la version 2 au lieu de la version 1. Cela réduit le nombre de renommages, ce qui améliore les performances de l'application. Cette modification n'affecte pas :
    - Les applications autres que Spark.
    - Applications qui écrivent sur d'autres systèmes de fichiers, tels que HDFS (qui utilise toujours la version 1 de FileOutputCommitter).
    - Les applications utilisant d'autres formats de sortie, tels que le texte ou le csv, qui utilisent déjà l'écriture directe EMRFS.

Problèmes connus

- JupyterHub
  - L'utilisation de classifications de configuration pour configurer JupyterHub des blocs-notes Jupyter individuels lors de la création d'un cluster n'est pas prise en charge. Modifiez manuellement le fichier jupyterhub_config.py et les fichiers jupyter_notebook_config.py

pour chaque utilisateur. Pour de plus amples informations, veuillez consulter Configuration JupyterHub.

- JupyterHub ne démarre pas sur des clusters au sein d'un sous-réseau privé, ce qui entraîne un échec du message`Error: ENOENT: no such file or directory, open '/etc/jupyter/conf/server.crt'` . Ce problème est dû à une erreur dans le script qui génère des certificats auto-signés. Utilisez la solution suivante pour générer des certificats auto-signés. Toutes les commandes sont exécutées lorsque vous êtes connecté au nœud primaire.

  1. Copiez le script de génération de certificat du conteneur vers le nœud primaire :

     ```
     sudo docker cp jupyterhub:/tmp/gen_self_signed_cert.sh ./
     ```

  2. Utilisez un éditeur de texte pour modifier la ligne 23 afin de changer le nom d'hôte public en nom d'hôte local, comme indiqué ci-dessous :

     ```
     local hostname=$(curl -s $EC2_METADATA_SERVICE_URI/local-hostname)
     ```

  3. Exécutez le script pour générer des certificats auto-signés :

     ```
     sudo bash ./gen_self_signed_cert.sh
     ```

  4. Déplacez les fichiers de certificat générés par le script vers le répertoire `/etc/jupyter/conf/` :

     ```
     sudo mv /tmp/server.crt /tmp/server.key /etc/jupyter/conf/
     ```

  Vous pouvez vérifier que `tail` le `jupyter.log` fichier JupyterHub a redémarré et qu'il renvoie un code de réponse 200. Par exemple :

  ```
  tail -f /var/log/jupyter/jupyter.log
  ```

  Vous devriez obtenir une réponse similaire à la suivante :

  ```
  # [I 2018-06-14 18:56:51.356 JupyterHub app:1581] JupyterHub is now running at
   https://:9443/
  # 19:01:51.359 - info: [ConfigProxy] 200 GET /api/routes
  ```

- Après le redémarrage du nœud principal ou le redémarrage du contrôleur d'instance, les CloudWatch métriques ne seront pas collectées et la fonctionnalité de dimensionnement

automatique ne sera pas disponible dans les versions 5.14.0, 5.15.0 ou 5.16.0 d'Amazon EMR. Ce problème est résolu dans Amazon EMR 5.17.0.

Version 5.13.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.13.0. Les modifications ont été apportées à la version 5.12.0.

Mises à niveau

- Mise à niveau de Spark vers la version 2.3.0
- Mise à niveau HBase vers la version 1.4.2
- Mise à niveau vers de Presto vers la version 0.194
- Mise à niveau AWS SDK pour Java vers la version 1.11.297

Modifications, améliorations et problèmes résolus

- Hive
  - Rétroportage de HIVE-15436. Hive améliorée pour ne APIs renvoyer que les vues.

Problèmes connus

- MXNet ne possède actuellement pas de bibliothèques OpenCV.

Version 5.12.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.12.2. Les modifications ont été apportées à la version 5.12.1.

Date de parution initiale : 29 août 2018

Modifications, améliorations et problèmes résolus

- Cette version corrige une vulnérabilité de sécurité potentielle.

Version 5.12.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.12.1. Les modifications ont été apportées à la version 5.12.0.

Date de parution initiale : 29 mars 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités potentielles.

Version 5.12.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.12.0. Les modifications ont été apportées à la version 5.11.1.

Mises à niveau

- AWS SDK pour Java 1.11.238 ⇒ 1.11.267. Pour plus d'informations, consultez le AWS SDK for Java Change Log GitHub on.
- Hadoop 2.7.3 ⇒ 2.8.3. Pour plus d'informations, consultez Versions d'Apache Hadoop.
- Flink 1.3.2 ⇒ 1.4.0. Pour plus d'informations, consultez Annonce de publication d'Apache Flink 1.4.0.
- HBase 1.3.1 ⇒ 1.4.0. Pour plus d'informations, consultez l'annonce HBase de sortie.
- Hue 4.0.1 ⇒ 4.1.0. Pour plus d'informations, veuillez consulter les Notes de mise à jour.
- MxNet 0,12,0 ⇒ 10,0. Pour plus d'informations, consultez la section MXNet Change Log on GitHub.
- Presto 0.187 ⇒ 0.188. Pour plus d'informations, veuillez consulter les Notes de mise à jour.

Modifications, améliorations et problèmes résolus

- Hadoop
  - La propriété `yarn.resourcemanager.decommissioning.timeout` a été remplacée par `yarn.resourcemanager.nodemanager-graceful-decommission-timeout-secs`. Vous pouvez utiliser cette propriété pour personnaliser la réduction de la taille du cluster. Pour plus d'informations, consultez Réduction de la taille des clusters dans le Guide de gestion Amazon EMR.

  - La CLI Hadoop a ajouté l'option `-d` à la commande `cp` (copy), qui spécifie la copie directe. Vous pouvez l'utiliser pour éviter de créer un fichier `.COPYING` intermédiaire, ce qui accélère la copie de données entre Amazon S3. Pour plus d'informations, consultez HADOOP-12384.
- Pig
  - Ajout de la classification de configuration `pig-env`, qui simplifie la configuration des propriétés de l'environnement Pig. Pour de plus amples informations, veuillez consulter Configuration des applications.
- Presto
  - Ajout de la classification de configuration, `presto-connector-redshift` que vous pouvez utiliser pour configurer des valeurs dans le fichier de configuration `redshift.properties` de Presto. Pour plus d'informations, consultez Connecteur Redshift dans la documentation Presto, et Configuration des applications.
  - La prise en charge de Presto pour EMRFS a été ajoutée et constitue la configuration par défaut. Les versions précédentes d'Amazon EMR utilisaient PrestOS3FileSystem, qui était la seule option. Pour de plus amples informations, veuillez consulter Configuration d'EMRFS et de PrestOS3 FileSystem.

    > **Note**
    >
    > Si vous interrogez des données sous-jacentes dans Amazon S3 avec Amazon EMR version 5.12.0, des erreurs Presto peuvent se produire. Cela est dû au fait que Presto ne parvient pas à récupérer les valeurs de classification de configuration depuis `emrfs-site.xml`. Pour contourner le problème, créez un sous-répertoire `emrfs` sous `usr/lib/presto/plugin/hive-hadoop2/` et un lien symbolique dans `usr/lib/presto/plugin/hive-hadoop2/emrfs` vers le fichier `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` existant. Redémarrez ensuite le processus presto-server (`sudo presto-server stop` suivi de `sudo presto-server start`).
- Spark
  - SPARK-22036 rétroporté : la BigDecimal multiplication renvoie parfois la valeur nulle.

Problèmes connus

- MXNet n'inclut pas les bibliothèques OpenCV.

- SparkR n'est pas disponible pour les clusters créés à l'aide d'une AMI personnalisée car R n'est pas installé par défaut sur les nœuds du cluster.

Version 5.11.3

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.11.3. Les modifications ont été apportées à la version 5.11.2.

Date de parution initiale : 18 juillet 2019

Modifications, améliorations et problèmes résolus

- Mise à jour de l'AMI Amazon Linux par défaut pour Amazon EMR afin d'inclure d'importantes mises à jour de sécurité du noyau Linux, notamment le problème de déni de service TCP SACK (AWS-2019-005).

Version 5.11.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.11.2. Les modifications ont été apportées à la version 5.11.1.

Date de parution initiale : 29 août 2018

Modifications, améliorations et problèmes résolus

- Cette version corrige une vulnérabilité de sécurité potentielle.

Version 5.11.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.11.1. Il s'agit des modifications apportées à la version Amazon EMR 5.11.0.

Date de parution initiale : 22 janvier 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités associées à l'exécution spéculative (CVE-2017-5715, CVE-2017-5753 et CVE-2017-5754). Pour de plus amples informations, veuillez consulter https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Problèmes connus

- MXNet n'inclut pas les bibliothèques OpenCV.
- Hive 2.3.2 définit `hive.compute.query.using.stats=true` par défaut. Cela entraîne des requêtes pour obtenir des données à partir de statistiques existantes plutôt que directement à partir des données, ce qui peut être déroutant. Par exemple, si vous avez une table avec `hive.compute.query.using.stats=true` et que vous téléchargez de nouveaux fichiers vers la table `LOCATION`, l'exécution d'une demande `SELECT COUNT(*)` sur la table renvoie le nombre des statistiques, plutôt que de récupérer les lignes ajoutées.

  Pour contourner ce problème, utilisez la commande `ANALYZE TABLE` pour collecter de nouvelles statistiques, ou définissez `hive.compute.query.using.stats=false`. Pour en savoir plus, consultez Statistiques dans Hive dans la documentation Apache Hive.

Version 5.11.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.11.0. Il s'agit des modifications apportées à la version Amazon EMR 5.10.0.

Mises à niveau

Les applications et les composants suivants ont été mis à niveau dans cette édition pour inclure les versions suivantes :

- Hive 2.3.2
- Spark 2.2.1
- Kit SDK pour Java 1.11.238

Nouvelles fonctionnalités

- Spark
  - Ajout du paramètre `spark.decommissioning.timeout.threshold`, qui améliore le comportement de mise hors service de Spark lors de l'utilisation d'instances Spot. Pour de plus amples informations, veuillez consulter Configuration du comportement de mise hors service du nœud.
  - Ajout du `aws-sagemaker-spark-sdk` composant à Spark, qui installe Amazon SageMaker Spark et les dépendances associées pour l'intégration de Spark à Amazon SageMaker. Vous

pouvez utiliser Amazon SageMaker Spark pour créer des pipelines d'apprentissage automatique (ML) Spark à l'aide d'Amazon SageMaker Stages. Pour plus d'informations, consultez le fichier readme de SageMaker Spark sur GitHub et l'utilisation d'Apache Spark avec Amazon SageMaker dans le manuel Amazon SageMaker Developer Guide.

Problèmes connus

- MXNet n'inclut pas les bibliothèques OpenCV.
- Hive 2.3.2 définit `hive.compute.query.using.stats=true` par défaut. Cela entraîne des requêtes pour obtenir des données à partir de statistiques existantes plutôt que directement à partir des données, ce qui peut être déroutant. Par exemple, si vous avez une table avec `hive.compute.query.using.stats=true` et que vous téléchargez de nouveaux fichiers vers la table `LOCATION`, l'exécution d'une demande `SELECT COUNT(*)` sur la table renvoie le nombre des statistiques, plutôt que de récupérer les lignes ajoutées.

  Pour contourner ce problème, utilisez la commande `ANALYZE TABLE` pour collecter de nouvelles statistiques, ou définissez `hive.compute.query.using.stats=false`. Pour en savoir plus, consultez Statistiques dans Hive dans la documentation Apache Hive.

Version 5.10.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.10.0. Il s'agit des modifications apportées à la version Amazon EMR 5.9.0.

Mises à niveau

Les applications et les composants suivants ont été mis à niveau dans cette édition pour inclure les versions suivantes :

- AWS SDK pour Java 1,11,221
- Hive 2.3.1
- Presto 0.187

Nouvelles fonctionnalités

- Ajout de la prise en charge de l'authentification Kerberos. Pour plus d'informations, consultez Utilisation de l'authentification Kerberos dans le Guide de gestion Amazon EMR

- Ajout de la prise en charge des rôles IAM pour les demandes EMRFS à Amazon S3. Pour plus d'informations, consultez la section Configuration des rôles IAM pour les demandes EMRFS vers Amazon S3 dans le Guide de gestion Amazon EMR.
- Ajout de la prise en charge des types d'instances P2 et P3 basées sur GPU. Pour plus d'informations, consultez les sections Instances Amazon EC2 P2 et instances Amazon EC2 P3. Les pilotes NVIDIA 384.81 et CUDA 9.0.176 sont installés sur ces types d'instance par défaut.
- Ajout de la prise en charge de Apache MXNet.

Modifications, améliorations et problèmes résolus

- Presto
  - Ajout de la prise en charge de l'utilisation du catalogue de données AWS Glue comme métastore Hive par défaut. Pour plus d'informations, consultez la section Utilisation de Presto avec le catalogue de données AWS Glue.
  - Ajout de la prise en charge des fonctions géospatiales.
  - Ajout de la prise en charge du déversement sur le disque pour les jointures.
  - Ajout de la prise en charge du connecteur Redshift.
- Spark
  - Rétroportage de SPARK-20640, ce qui permet de configurer le délai des appels de procédure distante et les tentatives aléatoires des enregistrements de réorganisation à l'aide des propriétés `spark.shuffle.registration.timeout` et `spark.shuffle.registration.maxAttempts`.
  - SPARK-21549 a été rétroporté, qui corrige une erreur survenue lors de l'écriture personnalisée OutputFormat dans des emplacements autres que le HDFS.
- Rétroportage de Hadoop-13270
- Les bibliothèques Numpy, Scipy et Matplotlib ont été supprimées de l'AMI Amazon EMR de base. Si vous avez besoin de ces bibliothèques pour votre application, elles sont disponibles dans le référentiel d'applications. Vous pouvez utiliser une action d'amorçage pour les installer sur tous les nœuds à l'aide de `yum install`.
- L'image AMI Amazon EMR de base n'inclut plus de packages RPM d'application, de telle manière que les packages RPM ne sont plus présents sur les nœuds du cluster. Custom AMIs et l'AMI de base Amazon EMR font désormais référence au référentiel de packages RPM dans Amazon S3.
- En raison de l'introduction de la facturation à la seconde sur Amazon EC2, le comportement de réduction par défaut est désormais de terminer la tâche à la fin de la tâche plutôt que de terminer

à l'heure de l'instance. Pour plus d'informations, consultez Configuration de la réduction de la capacité des clusters.

Problèmes connus

- MXNet n'inclut pas les bibliothèques OpenCV.
- Hive 2.3.1 définit `hive.compute.query.using.stats=true` par défaut. Cela entraîne des requêtes pour obtenir des données à partir de statistiques existantes plutôt que directement à partir des données, ce qui peut être déroutant. Par exemple, si vous avez une table avec `hive.compute.query.using.stats=true` et que vous téléchargez de nouveaux fichiers vers la table `LOCATION`, l'exécution d'une demande `SELECT COUNT(*)` sur la table renvoie le nombre des statistiques, plutôt que de récupérer les lignes ajoutées.

  Pour contourner ce problème, utilisez la commande `ANALYZE TABLE` pour collecter de nouvelles statistiques, ou définissez `hive.compute.query.using.stats=false`. Pour en savoir plus, consultez Statistiques dans Hive dans la documentation Apache Hive.

Version 5.9.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.9.0. Il s'agit des modifications apportées à la version Amazon EMR 5.8.0.

Date de parution : 5 octobre 2017

Dernière date de mise à jour de la fonctionnalité : 12 octobre 2017

Mises à niveau

Les applications et les composants suivants ont été mis à niveau dans cette édition pour inclure les versions suivantes :

- AWS SDK pour Java version 1.11.183
- Flink 1.3.2
- Hue 4.0.1
- Pig 0.17.0
- Presto 0.184

Nouvelles fonctionnalités

- Ajout de la prise en charge de Livy (version 0.4.0-incubating). Pour de plus amples informations, veuillez consulter Apache Livy.
- Ajout de la prise en charge de Hue Notebook pour Spark.
- Ajout de la prise en charge des EC2 instances Amazon de la série i3 (12 octobre 2017).

Modifications, améliorations et problèmes résolus

- Spark
  - Ajout d'un nouvel ensemble de fonctionnalités qui permet de s'assurer que Spark gère plus élégamment la terminaison des nœuds suite à un redimensionnement manuel ou à une demande de stratégie de dimensionnement automatique. Pour de plus amples informations, veuillez consulter Configuration du comportement de mise hors service du nœud.
  - Le protocole SSL est utilisé à la place du 3DES pour le chiffrement en transit du service de transfert par blocs, ce qui améliore les performances lors de l'utilisation de types d' EC2 instances Amazon avec AES-NI.
  - Rétroportage de SPARK-21494.
- Zeppelin
  - Rétroportage de ZEPPELIN-2377.
- HBase
  - Ajout du correctif HBASE-18533, qui autorise des valeurs supplémentaires pour la HBase BucketCache configuration en utilisant la classification de configuration. `hbase-site`
- Hue
  - Ajout de la prise en charge du catalogue de données AWS Glue pour l'éditeur de requêtes Hive dans Hue.
  - Par défaut, les superutilisateurs de Hue peuvent accéder à tous les fichiers auxquels les rôles IAM Amazon EMR sont autorisés à accéder. Les utilisateurs nouvellement créés n'ont pas automatiquement les autorisations d'accéder à l'explorateur de fichiers Amazon S3 et doivent avoir les autorisations `filebrowser.s3_access` activées pour leur groupe.
- Résolution du problème qui entraînait l'impossibilité d'accéder aux données JSON sous-jacentes créées avec le catalogue de données AWS Glue.

Problèmes connus

- Le lancement du cluster échoue lorsque toutes les applications sont installées et que la taille du volume racine Amazon EBS par défaut n'est pas modifiée. Pour contourner le problème, utilisez la `aws emr create-cluster` commande du AWS CLI et spécifiez un `--ebs-root-volume-size` paramètre plus grand.
- Hive 2.3.0 définit `hive.compute.query.using.stats=true` par défaut. Cela entraîne des requêtes pour obtenir des données à partir de statistiques existantes plutôt que directement à partir des données, ce qui peut être déroutant. Par exemple, si vous avez une table avec `hive.compute.query.using.stats=true` et que vous téléchargez de nouveaux fichiers vers la table `LOCATION`, l'exécution d'une demande `SELECT COUNT(*)` sur la table renvoie le nombre des statistiques, plutôt que de récupérer les lignes ajoutées.

  Pour contourner ce problème, utilisez la commande `ANALYZE TABLE` pour collecter de nouvelles statistiques, ou définissez `hive.compute.query.using.stats=false`. Pour en savoir plus, consultez Statistiques dans Hive dans la documentation Apache Hive.

Version 5.8.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.8.2. Les modifications ont été apportées à la version 5.8.1.

Date de parution initiale : 29 mars 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités potentielles.

Version 5.8.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.8.1. Il s'agit des modifications apportées à la version Amazon EMR 5.8.0.

Date de parution initiale : 22 janvier 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités associées à l'exécution spéculative (CVE-2017-5715, CVE-2017-5753 et CVE-2017-5754). Pour de plus amples informations, veuillez consulter https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Version 5.8.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.8.0. Il s'agit des modifications apportées à la version Amazon EMR 5.7.0.

Date de parution initiale : 10 août 2017

Dernière date de mise à jour de la fonctionnalité : 25 septembre 2017

Mises à niveau

Les applications et les composants suivants ont été mis à jour dans cette version pour inclure les versions suivantes :

- AWS SDK 1.11.160
- Flink 1.3.1
- Hive 2.3.0. Pour plus d'informations, consultez Notes de mise à jour sur le site Apache Hive.
- Spark 2.2.0. Pour plus d'informations, consultez Notes de mise à jour sur le site Apache Spark.

Nouvelles fonctionnalités

- Ajout de la prise en charge de l'affichage de l'historique de l'application (25 septembre 2017). Pour plus d'informations, consultez la section Affichage de l'historique des applications dans le Guide de gestion Amazon EMR.

Modifications, améliorations et problèmes résolus

- Intégration avec AWS Glue Data Catalog
  - Hive et Spark SQL ont désormais la possibilité d'utiliser AWS Glue Data Catalog comme magasin de métadonnées Hive. Pour plus d’informations, consultez Utilisation du catalogue de

données AWS Glue comme métastore pour Hive et Utiliser le catalogue AWS Glue Data Catalog avec Spark sur Amazon EMR.

- Historique de l'application ajouté aux détails du cluster, ce qui vous permet d'afficher les données historiques des applications YARN et des détails supplémentaires pour les applications Spark. Pour plus d'informations, consultez Afficher l'historique de l'application dans le Guide de gestion Amazon EMR.
- Oozie
  - Rétroportage de OOZIE-2748.
- Hue
  - Rétroportage de HUE-5859
- HBase
  - Ajout d'un correctif pour exposer l'heure de démarrage du serveur HBase principal via Java Management Extensions (JMX) à l'aide `getMasterInitializedTime` de.
  - Ajout d'un correctif qui améliore la date de début du cluster.

Problèmes connus

- Le lancement du cluster échoue lorsque toutes les applications sont installées et que la taille du volume racine Amazon EBS par défaut n'est pas modifiée. Pour contourner le problème, utilisez la `aws emr create-cluster` commande du AWS CLI et spécifiez un `--ebs-root-volume-size` paramètre plus grand.
- Hive 2.3.0 définit `hive.compute.query.using.stats=true` par défaut. Cela entraîne des requêtes pour obtenir des données à partir de statistiques existantes plutôt que directement à partir des données, ce qui peut être déroutant. Par exemple, si vous avez une table avec `hive.compute.query.using.stats=true` et que vous téléchargez de nouveaux fichiers vers la table `LOCATION`, l'exécution d'une demande `SELECT COUNT(*)` sur la table renvoie le nombre des statistiques, plutôt que de récupérer les lignes ajoutées.

  Pour contourner ce problème, utilisez la commande `ANALYZE TABLE` pour collecter de nouvelles statistiques, ou définissez `hive.compute.query.using.stats=false`. Pour en savoir plus, consultez Statistiques dans Hive dans la documentation Apache Hive.
- Spark – Il existe un problème de fuite du gestionnaire de fichiers avec le démon apppusher lors de l'utilisation de Spark, qui peut apparaître après plusieurs heures ou jours pour une tâche Spark de longue durée. Pour corriger le problème, connectez-vous au nœud maître et tapez `sudo /`

`etc/init.d/apppusher stop`. Cette commande arrête le démon apppusher, qu'Amazon EMR redémarre automatiquement.

- Historique de l'application
  - Les données d'historique pour les exécuteurs Spark inactifs ne sont pas disponibles.
  - L'historique de l'application n'est pas disponible pour les clusters qui utilisent une configuration de sécurité pour permettre le chiffrement à la volée.

Version 5.7.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.7.0. Il s'agit des modifications apportées à la version Amazon EMR 5.6.0.

Date de parution : 13 juillet 2017

Mises à niveau

- Flink 1.3.0
- Phoenix 4.11.0
- Zeppelin 0.7.2

Nouvelles fonctionnalités

- Possibilité ajoutée de spécifier une AMI Amazon Linux personnalisée lorsque vous créez un cluster. Pour en savoir plus, consultez Utilisation d'une image AMI personnalisée.

Modifications, améliorations et problèmes résolus

- HBase
  - Ajout de la possibilité de configurer des clusters HBase en lecture et réplication. Consultez Utilisation d'un cluster de réplica en lecture.
  - Correctifs de nombreux bogues et améliorations
- Presto – ajout de la possibilité de configurer `node.properties`.
- YARN – ajout de la possibilité de configurer `container-log4j.properties`.
- Sqoop – SQOOP-2880 rétroporté, qui introduit un argument vous permettant de définir le répertoire temporaire de Sqoop.

Version 5.6.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.6.0. Il s'agit des modifications apportées à la version Amazon EMR 5.5.0.

Date de parution : 5 juin 2017

Mises à niveau

- Flink 1.2.1
- HBase 1.3.1
- Mahout 0.13.0. Il s'agit de la première version de Mahout à prendre en charge Spark 2.x dans les versions 5.0 et ultérieures d'Amazon EMR.
- Spark 2.1.1

Modifications, améliorations et problèmes résolus

- Presto
  - Ajout de la possibilité d'activer une communication SSL/TLS sécurisée entre les nœuds Presto en activant le chiffrement en transit à l'aide d'une configuration de sécurité. Pour plus d'informations, consultez Chiffrement des données en transit.
  - Rétroportage de Presto 7661, ce qui ajoute l'option `VERBOSE` à l'instruction `EXPLAIN ANALYZE` pour transmettre des statistiques de bas niveau plus détaillées sur un plan de la requête.

Version 5.5.3

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.5.3. Les modifications ont été apportées à la version 5.5.2.

Date de parution initiale : 29 août 2018

Modifications, améliorations et problèmes résolus

- Cette version corrige une vulnérabilité de sécurité potentielle.

Version 5.5.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.5.2. Les modifications ont été apportées à la version 5.5.1.

Date de parution initiale : 29 mars 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités potentielles.

Version 5.5.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.5.1. Il s'agit des modifications apportées à la version Amazon EMR 5.5.0.

Date de parution initiale : 22 janvier 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités associées à l'exécution spéculative (CVE-2017-5715, CVE-2017-5753 et CVE-2017-5754). Pour de plus amples informations, veuillez consulter https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Version 5.5.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.5.0. Il s'agit des modifications apportées à la version Amazon EMR 5.4.0.

Date de parution : 26 avril 2017

Mises à niveau

- Hue 3.12
- Presto 0.170
- Zeppelin 0.7.1

- ZooKeeper 3.4.10

Modifications, améliorations et problèmes résolus

- Spark
  - DAGScheduler Correctif de Spark rétroporté (SPARK-20115) permettant de recalculer tous les blocs de shuffle perdus lorsque le service de shuffle externe n'est pas disponible dans la version 2.1.0 de Spark, incluse dans cette version.
- Flink
  - Flink s'appuie désormais sur Scala 2.11. Si vous utilisez l'API et les bibliothèques Scala, nous vous recommandons d'utiliser Scala 2.11 dans vos projets.
  - Résolution d'un problème où les valeurs par défaut `HADOOP_CONF_DIR` et `YARN_CONF_DIR` n'étaient pas définies correctement, si bien que `start-scala-shell.sh` ne fonctionnait pas. Ajout également de la possibilité de définir ces valeurs à l'aide de `env.hadoop.conf.dir` et `env.yarn.conf.dir` dans `/etc/flink/conf/flink-conf.yaml` ou la classification de configuration `flink-conf`.
  - Introduction d'une nouvelle commande spécifique à EMR, `flink-scala-shell` comme wrapper pour `start-scala-shell.sh`. Nous vous recommandons d'utiliser cette commande au lieu de `start-scala-shell`. La nouvelle commande simplifie l'exécution. Par exemple, `flink-scala-shell -n 2` démarre un shell Flink Scala avec un parallélisme de 2 tâches.
  - Introduction d'une nouvelle commande spécifique à EMR, `flink-yarn-session` comme wrapper pour `yarn-session.sh`. Nous vous recommandons d'utiliser cette commande au lieu de `yarn-session`. La nouvelle commande simplifie l'exécution. Par exemple, `flink-yarn-session -d -n 2` démarre une session Flink de longue durée à l'état détaché avec deux gestionnaires de tâches.
  - Résolution du problème (FLINK-6125) commons httpclient n'est plus nuancé dans Flink 1.2.
- Presto
  - Ajout de la prise en charge de l'authentification LDAP. L'utilisation de LDAP avec Presto sur Amazon EMR nécessite que vous activiez l'accès HTTPS pour le coordinateur Presto (`http-server.https.enabled=true` dans `config.properties`). Pour plus de détails sur la configuration, consultez Authentification LDAP dans la documentation Presto.
  - Ajout de la prise en charge de `SHOW GRANTS`.
- AMI Linux Amazon EMR de base

- Les versions d'Amazon EMR sont désormais basées sur Amazon Linux 2017.03. Pour plus d'informations, consultez Notes de mise à jour Amazon Linux AMI 2017.03.
- Suppression de Python 2.6 de l'image Linux de base d'Amazon EMR. Python 2.7 et 3.4 sont installés par défaut. Vous pouvez installer Python 2.6 manuellement si nécessaire.

Version 5.4.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.4.0. Il s'agit des modifications apportées à la version Amazon EMR 5.3.0.

Date de parution : 8 mars 2017

Mises à niveau

Les mises à niveau suivantes sont disponibles dans cette version :

- Mise à niveau vers Flink 1.2.0
- Mise à niveau vers HBase 1.3.0
- Mise à niveau vers Phoenix 4.9.0

> **Note**
>
> Si vous effectuez une mise à niveau depuis une version antérieure d'Amazon EMR vers Amazon EMR version 5.4.0 ou supérieure et utilisez une indexation secondaire, mettez à niveau les index locaux comme décrit dans la documentation Apache Phoenix. Amazon EMR supprime les configurations requises de la classification `hbase-site`, mais les index doivent être repeuplés. La mise à niveau en ligne et hors ligne des index est prise en charge. Les mises à niveau sont par défaut en ligne, ce qui signifie que les index sont remplis de nouveau lors de l'initialisation depuis les clients Phoenix en version 4.8.0 ou supérieure. Pour spécifier des mises à niveau hors ligne, définissez la configuration `phoenix.client.localIndexUpgrade` sur false dans la classification `phoenix-site`, puis lancez SSH sur le nœud maître afin d'exécuter `psql [zookeeper] -1`.

- Mise à niveau vers Presto 0.166
- Mise à niveau vers Zeppelin 0.7.0

Modifications et améliorations

Les modifications suivantes sont apportées aux versions Amazon EMR pour l'étiquette de version emr-5.4.0 :

- Ajout de la prise en charge des instances r4. Consultez la section Types d' EC2 instances Amazon.

Version 5.3.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.3.1. Il s'agit des modifications apportées à la version Amazon EMR 5.3.0.

Date de version : 7 février 2017

Modifications mineures pour rétroporter les correctifs Zeppelin et mettre à jour l'AMI par défaut pour Amazon EMR.

Version 5.3.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.3.0. Il s'agit des modifications apportées à la version Amazon EMR 5.2.1.

Date de parution : 26 janvier 2017

Mises à niveau

Les mises à niveau suivantes sont disponibles dans cette version :

- Mise à niveau vers Hive 2.1.1
- Mise à niveau vers Hue 3.11.0
- Mise à niveau vers Spark 2.1.0
- Mise à niveau vers Oozie 4.3.0
- Mise à niveau vers Flink 1.1.4

Modifications et améliorations

Les modifications suivantes sont apportées aux versions Amazon EMR pour l'étiquette de version emr-5.3.0 :

- Ajout à Hue d'un correctif qui vous permet d'utiliser le paramètre `interpreters_shown_on_wheel` pour configurer les interpréteurs à afficher en premier sur la roue de sélection Notebook quel que soit leur ordre dans le fichier `hue.ini`.
- Ajout de la classification de configuration, `hive-parquet-logging` que vous pouvez utiliser pour configurer des valeurs dans le fichier `parquet-logging.properties` de Hive.

Version 5.2.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.2.2. Il s'agit des modifications apportées à la version Amazon EMR 5.2.1.

Date de parution : 2 mai 2017

Problèmes connus résolus depuis les versions précédentes

- SPARK-194459 a été rétroporté, ce qui résout un problème en cas d'échec de la lecture d'une table ORC contenant des colonnes. char/varchar

Version 5.2.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.2.1. Il s'agit des modifications apportées à la version Amazon EMR 5.2.0.

Date de parution : 29 décembre 2016

Mises à niveau

Les mises à niveau suivantes sont disponibles dans cette version :

- Mise à niveau vers Presto 0.157.1. Pour plus d'informations, consultez Notes de mise à jour Presto dans la documentation Presto.
- Mise à niveau vers Zookeeper 3.4.9. Pour plus d'informations, consultez les notes ZooKeeper de publication dans la ZooKeeper documentation d'Apache.

Modifications et améliorations

Les modifications suivantes sont apportées aux versions Amazon EMR pour l'étiquette de version emr-5.2.1 :

- Ajout de la prise en charge du type d' EC2 instance Amazon m4.16xlarge dans Amazon EMR version 4.8.3 et versions ultérieures, à l'exception des versions 5.0.0, 5.0.3 et 5.2.0.
- Les versions d'Amazon EMR sont désormais basées sur Amazon Linux 2016.09. Pour de plus amples informations, veuillez consulter https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.
- L'emplacement des chemins de configuration Flink et YARN sont désormais définis par défaut dans `/etc/default/flink`, et vous n'avez pas besoin de définir les variables d'environnement `FLINK_CONF_DIR` et `HADOOP_CONF_DIR` lorsque vous exécutez les scripts de pilote `flink` ou `yarn-session.sh` pour lancer des tâches Flink.
- Ajout du support pour les FlinkKinesisConsumer cours.

Problèmes connus résolus depuis les versions précédentes

- Correction d'un problème dans Hadoop où le ReplicationMonitor thread pouvait rester bloqué pendant longtemps en raison d'une course entre la réplication et la suppression du même fichier dans un grand cluster.
- Correction d'un problème en raison duquel ControlledJob #toString échouait avec une exception de pointeur nul (NPE) lorsque le statut de la tâche n'était pas correctement mis à jour.

Version 5.2.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.2.0. Il s'agit des modifications apportées à la version Amazon EMR 5.1.0.

Date de parution : 21 novembre 2016

Modifications et améliorations

Les modifications et les améliorations suivantes sont disponibles dans cette version :

- Ajout du mode de stockage Amazon S3 pour HBase.
- Vous permet de spécifier un emplacement Amazon S3 pour le HBase rootdir. Pour plus d'informations, consultez HBaseAmazon S3.

Mises à niveau

Les mises à niveau suivantes sont disponibles dans cette version :

- Mise à niveau vers Spark 2.0.2

Problèmes connus résolus depuis les versions précédentes

- Correction d'un problème avec /mnt tout en étant limité à 2 To sur les types d'instance uniquement EBS.
- Correction d'un problème avec les journaux instance-controller et logpusher édités dans leurs fichiers .out correspondants au lieu de leurs fichiers .log configurés log4j, qui permutent toutes les heures. Les fichiers .out n'effectuent pas de permutation, cela remplirait finalement la partition /emr. Ce problème affecte uniquement les types d'instance de virtualisation HVM.

Version 5.1.0

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.1.0. Il s'agit des modifications apportées à la version Amazon EMR 5.0.0.

Date de parution : 3 novembre 2016

Modifications et améliorations

Les modifications et les améliorations suivantes sont disponibles dans cette version :

- Support ajouté pour Flink 1.1.3.
- Presto a été ajouté comme une option dans la section ordinateur portable de Hue.

Mises à niveau

Les mises à niveau suivantes sont disponibles dans cette version :

- Mise à niveau vers la version HBase 1.2.3
- Mise à niveau vers Zeppelin 0.6.2

Problèmes connus résolus depuis les versions précédentes

- Correction d'un problème avec des requêtes Tez sur Amazon S3 où les fichiers ORC ne fonctionnaient pas aussi bien qu'avec les versions antérieures d'Amazon EMR 4.x.

Version 5.0.3

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 5.0.3. Il s'agit des modifications apportées à la version Amazon EMR 5.0.0.

Date de parution : 24 octobre 2016

Mises à niveau

Les mises à niveau suivantes sont disponibles dans cette version :

- Mise à niveau vers Hadoop 2.7.3
- Mise à niveau vers Presto 0.152.3, qui comprend la prise en charge de l'interface Web Presto. Utilisez le port 8889 du coordinateur Presto pour accéder à l'interface Web Presto. Pour plus d'informations sur l'interface Web Presto, consultez Interface Web dans la documentation Presto.
- Mise à niveau vers Spark 2.0.1
- Les versions d'Amazon EMR sont désormais basées sur Amazon Linux 2016.09. Pour de plus amples informations, veuillez consulter https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Version 5.0.0

Date de parution : 27 juillet 2016

Mises à niveau

Les mises à niveau suivantes sont disponibles dans cette version :

- Mise à niveau vers Hive 2.1
- Mise à niveau vers Presto 0.150
- Mise à niveau vers Spark 2.0
- Mise à niveau vers Hue 3.10.0
- Mise à niveau vers Pig 0.16.0
- Mise à niveau vers Tez 0.8.4
- Mise à niveau vers Zeppelin 0.6.1

Modifications et améliorations

Les modifications suivantes sont apportées aux versions Amazon EMR pour l'étiquette de version emr-5.0.0 ou ultérieures :

- Amazon EMR prend en charge les dernières versions open source de Hive (version 2.1) et de Pig (version 0.16.0). Si vous utilisiez Hive ou Pig sur Amazon EMR auparavant, cela risque d'avoir un impact sur certains cas d'utilisation. Pour en savoir plus, consultez Hive et Pig.
- Tez est désormais le moteur d'exécution par défaut de Hive et Pig. Pour changer cela, vous devez modifier les valeurs appropriées dans les classifications de configuration `hive-site` et `pig-properties`, respectivement.
- Une fonction améliorée de débogage des étapes a été ajoutée, ce qui vous permet de voir la cause première des échecs des étapes si le service peut déterminer la cause. Pour plus d'informations, consultez la section Débogage d'étape amélioré dans le Guide de gestion Amazon EMR.
- Les applications qui se terminaient par « -Sandbox » n'ont plus ce suffixe. Cela peut interrompre l'automatisation de certaines tâches, par exemple, si vous utilisez des scripts pour lancer des clusters avec ces applications. Le tableau suivant montre les noms d'application dans Amazon EMR 4.7.2 contre Amazon EMR 5.0.0.

  Modification des noms d'application

Amazon EMR 4.7.2	Amazon EMR 5.0.0
Oozie-Sandbox	Oozie
Presto-Sandbox	Presto
Sqoop-Sandbox	Sqoop
Zeppelin-Sandbox	Zeppelin
ZooKeeper-Bac à sable	ZooKeeper

- Spark est maintenant compilé pour Scala 2.11.
- Java 8 est désormais la JVM par défaut. Toutes les applications s'exécutent avec le runtime Java 8. Aucune modification n'est apportée bytecode cible des applications. La plupart des applications continuent de cibler Java 7.
- Zeppelin inclut désormais des fonctions d'authentification. Pour en savoir plus, consultez Zeppelin.

- Ajout de la prise en charge des configurations de sécurité, qui vous permettent de créer et d'appliquer des options de chiffrement plus facilement. Pour en savoir plus, consultez Chiffrement des données.

Version 4.9.5

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.9.5. Les modifications ont été apportées à la version 4.9.4.

Date de parution initiale : 29 août 2018

Modifications, améliorations et problèmes résolus

- HBase
  - Cette version corrige une vulnérabilité de sécurité potentielle.

Version 4.9.4

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.9.4. Les modifications ont été apportées à la version 4.9.3.

Date de parution initiale : 29 mars 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités potentielles.

Version 4.9.3

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.9.3. Il s'agit des modifications apportées à la version Amazon EMR 4.9.2.

Date de parution initiale : 22 janvier 2018

Modifications, améliorations et problèmes résolus

- Mise à jour du noyau Amazon Linux de l'AMI Amazon Linux par défaut pour Amazon EMR afin de corriger les vulnérabilités associées à l'exécution spéculative (CVE-2017-5715, CVE-2017-5753 et

CVE-2017-5754). Pour de plus amples informations, veuillez consulter https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Version 4.9.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.9.2. Il s'agit des modifications apportées à la version Amazon EMR 4.9.1.

Date de parution : 13 juillet 2017

Des modifications mineures, des correctifs de bogues et des améliorations ont été apportées à cette version.

Version 4.9.1

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.9.1. Il s'agit des modifications apportées à la version Amazon EMR 4.8.4.

Date de parution : 10 avril 2017

Problèmes connus résolus depuis les versions précédentes

- Rétroportages de HIVE-9976 et HIVE-10106
- Résolution d'un problème dans YARN où un nombre important de nœuds (plus de 2 000) et de conteneurs (plus de 5 000) provoquait une erreur de mémoire insuffisante, par exemple : `"Exception in thread 'main' java.lang.OutOfMemoryError"`.

Modifications et améliorations

Les modifications suivantes sont apportées aux versions Amazon EMR pour l'étiquette de version emr-4.9.1 :

- Les versions d'Amazon EMR sont désormais basées sur Amazon Linux 2017.03. Pour de plus amples informations, veuillez consulter https://aws.amazon.com/amazon-linux-ami/2017.03-release-notes/.
- Suppression de Python 2.6 de l'image Linux de base d'Amazon EMR. Vous pouvez installer Python 2.6 manuellement si nécessaire.

Version 4.8.4

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.8.4. Il s'agit des modifications apportées à la version Amazon EMR 4.8.3.

Date de parution : 7 février 2017

Des modifications mineures, des correctifs de bogues et des améliorations ont été apportées à cette version.

Version 4.8.3

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.8.3. Il s'agit des modifications apportées à la version Amazon EMR 4.8.2.

Date de parution : 29 décembre 2016

Mises à niveau

Les mises à niveau suivantes sont disponibles dans cette version :

- Mise à niveau vers Presto 0.157.1. Pour plus d'informations, consultez Notes de mise à jour Presto dans la documentation Presto.
- Mise à niveau vers Spark 1.6.3. Pour plus d'informations, consultez les Notes de mise à jour Spark dans la documentation Apache Spark.
- Mise à niveau vers la version ZooKeeper 3.4.9. Pour plus d'informations, consultez les notes ZooKeeper de publication dans la ZooKeeper documentation d'Apache.

Modifications et améliorations

Les modifications suivantes sont apportées aux versions Amazon EMR pour l'étiquette de version emr-4.8.3 :

- Ajout de la prise en charge du type d' EC2 instance Amazon m4.16xlarge dans Amazon EMR version 4.8.3 et versions ultérieures, à l'exception des versions 5.0.0, 5.0.3 et 5.2.0.
- Les versions d'Amazon EMR sont désormais basées sur Amazon Linux 2016.09. Pour de plus amples informations, veuillez consulter https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Problèmes connus résolus depuis les versions précédentes

- Correction d'un problème dans Hadoop où le ReplicationMonitor thread pouvait rester bloqué pendant longtemps en raison d'une course entre la réplication et la suppression du même fichier dans un grand cluster.
- Correction d'un problème en raison duquel ControlledJob #toString échouait avec une exception de pointeur nul (NPE) lorsque le statut de la tâche n'était pas correctement mis à jour.

Version 4.8.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.8.2. Il s'agit des modifications apportées à la version Amazon EMR 4.8.0.

Date de parution : 24 octobre 2016

Mises à niveau

Les mises à niveau suivantes sont disponibles dans cette version :

- Mise à niveau vers Hadoop 2.7.3
- Mise à niveau vers Presto 0.152.3, qui comprend la prise en charge de l'interface Web Presto. Utilisez le port 8889 du coordinateur Presto pour accéder à l'interface Web Presto. Pour plus d'informations sur l'interface Web Presto, consultez Interface Web dans la documentation Presto.
- Les versions d'Amazon EMR sont désormais basées sur Amazon Linux 2016.09. Pour de plus amples informations, veuillez consulter https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Version 4.8.0

Date de parution : 7 septembre 2016

Mises à niveau

Les mises à niveau suivantes sont disponibles dans cette version :

- Mise à niveau vers la version HBase 1.2.2
- Mise à niveau vers Presto-Sandbox 0.151

- Mise à niveau vers Tez 0.8.4
- Mise à niveau vers Zeppelin-Sandbox 0.6.1

Modifications et améliorations

Les modifications suivantes sont apportées aux versions Amazon EMR pour l'étiquette de version emr-4.8.0 :

- Correction d'un problème dans YARN à cause duquel ils ApplicationMaster tentaient de nettoyer des conteneurs qui n'existent plus parce que leurs instances avaient été fermées.
- Correction de l'URL hive-server2 pour les actions Hive2 dans les exemples Oozie.
- Ajout de la prise en charge d'autres catalogues Presto.
- Correctifs rétroportés : HIVE-8948, HIVE-12679, HIVE-13405, PHOENIX-3116, HADOOP-12689
- Ajout de la prise en charge des configurations de sécurité, qui vous permettent de créer et d'appliquer des options de chiffrement plus facilement. Pour en savoir plus, consultez Chiffrement des données.

Version 4.7.2

Les notes de mises à jour suivantes incluent des informations sur la version Amazon EMR 4.7.2.

Date de parution : 15 juillet 2016

Fonctionnalités

Les fonctions suivantes sont disponibles dans cette version :

- Mise à niveau vers Mahout 0.12.2
- Mise à niveau vers Presto 0.148
- Mise à niveau vers Spark 1.6.2
- Vous pouvez désormais créer un AWSCredentials fournisseur à utiliser avec EMRFS en utilisant un URI comme paramètre. Pour plus d'informations, voir Création d'un AWSCredentials fournisseur pour EMRFS.
- EMRFS permet maintenant aux utilisateurs de configurer un point de terminaison DynamoDB personnalisé pour les métadonnées de leur vue cohérente à l'aide de la propriété `fs.s3.consistent.dynamodb.endpoint` dans le fichier `emrfs-site.xml`.

- Ajout d'un script dans `/usr/bin` appelé `spark-example`, qui encapsule `/usr/lib/spark/spark/bin/run-example` pour vous permettre d'exécuter des exemples directement. Par exemple, pour exécuter l' SparkPi exemple fourni avec la distribution Spark, vous pouvez l'exécuter `spark-example SparkPi 100` à partir de la ligne de commande ou en `command-runner.jar` tant qu'étape dans l'API.

Problèmes connus résolus depuis les versions précédentes

- Résolution d'un problème où `spark-assembly.jar` pour Oozie n'était pas à l'emplacement approprié quand Spark était également installé, ce qui provoquait un échec du lancement d'applications Spark avec Oozie.
- Résolution d'un problème lié à la journalisation basée sur Spark Log4j dans des conteneurs YARN.

Version 4.7.1

Date de parution : 10 juin 2016

Problèmes connus résolus depuis les versions précédentes

- Résolution d'un problème qui augmentait la durée de démarrage des clusters lancés dans un VPC avec des sous-réseaux privés. Ce bogue affectait uniquement les clusters lancés avec la version 4.7.0 d'Amazon EMR.
- Résolution d'un problème où des listes de fichiers dans Amazon EMR n'étaient pas traitées correctement pour les clusters lancés avec la version 4.7.0 d'Amazon EMR.

Version 4.7.0

> **Important**
>
> Amazon EMR 4.7.0 est obsolète. Veuillez plutôt utiliser Amazon EMR 4.7.1 ou une version ultérieure.

Date de parution : 2 juin 2016

Fonctionnalités

Les fonctions suivantes sont disponibles dans cette version :

- Ajout d'Apache Phoenix 4.7.0
- Ajout d'Apache Tez 0.8.3
- Mise à niveau vers la version HBase 1.2.1
- Mise à niveau vers Mahout 0.12.0
- Mise à niveau vers Presto 0.147
- Mise à niveau de la AWS SDK pour Java version 1.10.75
- L'indicateur final a été supprimé de la propriété `mapreduce.cluster.local.dir` dans `mapred-site.xml` pour permettre aux utilisateurs d'exécuter Pig en mode local.

Pilotes JDBC Amazon Redshift disponibles sur le cluster

Les pilotes JDBC Amazon Redshift sont maintenant inclus dans `/usr/share/aws/redshift/jdbc`. `/usr/share/aws/redshift/jdbc/RedshiftJDBC41.jar` est le pilote Amazon Redshift compatible avec JDBC 4.1 et `/usr/share/aws/redshift/jdbc/RedshiftJDBC4.jar` est le pilote Amazon Redshift compatible avec JDBC 4.0. Pour plus d'informations, consultez la section Configuration d'une connexion JDBC dans le Guide de gestion Amazon Redshift.

Java 8

Sauf pour Presto, OpenJDK 1.7 est le JDK par défaut utilisé pour toutes les applications. Cependant, OpenJDK 1.7 et 1.8 sont installés. Pour plus d'informations sur la configuration de `JAVA_HOME` pour les applications, consultez Configuration d'applications pour utiliser Java 8.

Problèmes connus résolus depuis les versions précédentes

- Résolution d'un problème de noyau qui affectait de manière significative les performances sur les volumes HDD à débit optimisé (ST1) EBS pour Amazon EMR dans emr-4.6.0.
- Résolution d'un problème où un cluster échouait si une zone de chiffrement HDFS était spécifiée sans que Hadoop ait été choisi comme application.
- Remplacement de la stratégie d'écriture HDFS par défaut `RoundRobin` par `AvailableSpaceVolumeChoosingPolicy`. Certains volumes n'ont pas été utilisés

correctement avec la RoundRobin configuration, ce qui a entraîné la défaillance de nœuds principaux et un HDFS peu fiable.

- Résolution d'un problème lié à l'interface de ligne de commande EMRFS, qui entraînait une exception lors de la création de la table de métadonnées DynamoDB par défaut pour des vues cohérentes.
- Résolution d'un problème de blocage dans EMRFS qui pouvait potentiellement se produire lors d'opérations de changement de nom et de copie en plusieurs parties.
- Correction d'un problème lié à EMRFS qui faisait en sorte que la CopyPart taille par défaut était de 5 Mo. La valeur par défaut est maintenant correctement définie sur 128 Mo.
- Résolution d'un problème lié à la configuration upstart dans Zeppelin qui pouvait vous empêcher potentiellement d'arrêter le service.
- Résolution d'un problème lié à Spark et Zeppelin qui vous empêchait d'utiliser le schéma d'URI `s3a://`, car `/usr/lib/hadoop/hadoop-aws.jar` n'était pas chargé correctement dans leur chemin de classe respectif.
- Rétroportage de HUE-2484.
- J'ai rétroporté un commit depuis Hue 3.9.0 (il n'existe pas de JIRA) pour corriger un problème avec l' HBase exemple de navigateur.
- Rétroportage de HIVE-9073.

Version 4.6.0

Date de parution : 21 avril 2016

Fonctionnalités

Les fonctions suivantes sont disponibles dans cette version :

- Ajouté HBase 1.2.0
- Ajout de Zookeeper-Sandbox 3.4.8
- Mise à niveau vers Presto-Sandbox 0.143
- Les versions Amazon EMR sont désormais basées sur Amazon Linux 2016.03.0. Pour de plus amples informations, veuillez consulter https://aws.amazon.com/amazon-linux-ami/2016.03-release-notes/.

Problème affectant les types de volume HDD à débit optimisé (ST1) EBS

Un problème dans les versions 4.2 et précédentes de noyau Linux affecte de manière significative les performances sur les volumes HDD à débit optimisé (ST1) EBS pour EMR. Cette version (emr-4.6.0) utilise la version de noyau 4.4.5 et est donc concernée. Par conséquent, nous vous recommandons de ne pas utiliser emr-4.6.0 si vous voulez utiliser des volumes EBS st1. Vous pouvez utiliser emr-4.5.0 ou des versions précédentes d'Amazon EMR avec ST1 sans incidence. En outre, nous fournirons le correctif avec les versions futures.

Valeurs par défaut de Python

Python 3.4 est maintenant installé par défaut, mais Python 2.7 reste la valeur système par défaut. Vous pouvez configurer Python 3.4 comme valeur par défaut du système à l'aide d'une action bootstrap ; vous pouvez utiliser l'API de configuration pour définir l'exportation de PYSPARK_PYTHON sur `/usr/bin/python3.4` dans la classification afin d'affecter la `spark-env` version de Python utilisée par. PySpark

Java 8

Sauf pour Presto, OpenJDK 1.7 est le JDK par défaut utilisé pour toutes les applications. Cependant, OpenJDK 1.7 et 1.8 sont installés. Pour plus d'informations sur la configuration de `JAVA_HOME` pour les applications, consultez Configuration d'applications pour utiliser Java 8.

Problèmes connus résolus depuis les versions précédentes

- Résolution d'un problème où le provisionnement d'application échouait parfois de manière aléatoire en raison d'un mot de passe généré.
- Auparavant, `mysqld` était installé sur tous les nœuds. Maintenant, il est uniquement installé sur l'instance principale et seulement si l'application choisie inclut `mysql-server` comme composant. Actuellement, les applications suivantes incluent le `mysql-server` composant : Hive HCatalog, Hue, Presto-Sandbox et Sqoop-Sandbox.
- Remplacement de la valeur par défaut 80 par 32 pour `yarn.scheduler.maximum-allocation-vcores`, ce qui résout un problème introduit dans emr-4.4.0 qui se produit principalement avec Spark lors de l'utilisation de l'option `maximizeResourceAllocation` dans un cluster dont le type d'instance principal est l'un des types d'instance large pour lesquels des cœurs virtuels YARN sont définis sur des valeurs supérieures à 32. Les types c4.8xlarge, cc2.8xlarge, hs1.8xlarge, i2.8xlarge, m2.4xlarge, r3.8xlarge, d2.8xlarge et m4.10xlarge étaient affectés par ce problème.

- s3-dist-cp utilise désormais EMRFS pour toutes les nominations Amazon S3 et n'effectue plus une copie intermédiaire dans un répertoire HDFS temporaire.
- Résolution d'un problème lié au traitement des exceptions pour le chiffrement côté client du chargement partitionné.
- Ajout d'une option pour permettre aux utilisateurs de modifier la classe de stockage Amazon S3. Par défaut, ce paramètre est `STANDARD`. La configuration de classification `emrfs-site` est `fs.s3.storageClass`, et les valeurs possibles sont `STANDARD`, `STANDARD_IA` et `REDUCED_REDUNDANCY`. Pour plus d'informations sur les classes de stockage, consultez la section Classes de stockage dans le Guide de l'utilisateur Amazon Simple Storage Service.

Version 4.5.0

Date de parution : 4 avril 2016

Fonctionnalités

Les fonctions suivantes sont disponibles dans cette version :

- Mise à niveau vers Spark 1.6.1
- Mise à niveau vers Hadoop 2.7.2
- Mise à niveau vers Presto 0.140
- Ajout de la AWS KMS prise en charge du chiffrement côté serveur Amazon S3.

Problèmes connus résolus depuis les versions précédentes

- Résolution d'un problème où les serveurs MySQL et Apache ne démarraient pas après le redémarrage d'un nœud.
- Résolution d'un problème où IMPORT ne fonctionnait pas correctement avec les tables non partitionnées stockées dans Amazon S3
- Résolution d'un problème lié à Presto où le répertoire intermédiaire devait être `/mnt/tmp` au lieu de `/tmp` lors de l'écriture dans des tables Hive.

Version 4.4.0

Date de parution : 14 mars 2016

Fonctionnalités

Les fonctions suivantes sont disponibles dans cette version :

- Ajouté HCatalog 1.0.0
- Ajout de Sqoop-Sandbox 1.4.6
- Mise à niveau vers Presto 0.136
- Mise à niveau vers Zeppelin 0.5.6
- Mise à niveau vers Mahout 0.11.1
- `dynamicResourceAllocation` activé par défaut.
- Ajout d'un tableau de toutes les classifications de configuration pour la version. Pour en savoir plus, consultez le tableau Classifications des configurations dans Configuration des applications.

Problèmes connus résolus depuis les versions précédentes

- Correction d'un problème à cause duquel le `maximizeResourceAllocation` paramètre ne réservait pas suffisamment de mémoire pour les ApplicationMaster démons YARN.
- Résolution d'un problème rencontré avec un DNS personnalisé. Si des entrées dans `resolve.conf` précèdent les entrées personnalisées fournies, les entrées personnalisées ne peuvent pas être résolues. Ce comportement était affecté par des clusters dans un VPC dans lequel le serveur de noms VPC par défaut était inséré comme première entrée dans `resolve.conf`.
- Résolution d'un problème où la version Python par défaut passait à la version 2.7 et boto n'était pas installé pour cette version.
- Résolution d'un problème où des conteneurs YARN et des applications Spark généraient un fichier rdd (round robin database) unique Ganglia, si bien que le premier disque attaché à l'instance se remplissait. En raison de ce correctif, les métriques de niveau conteneur YARN ont été désactivées et les métriques de niveau application Spark ont été désactivées.
- Résolution d'un problème dans le transmetteur de journaux où celui-ci supprimait tous les dossiers de journal vides. De ce fait, l'interface de ligne de commande Hive ne pouvait pas journaliser, car le transmetteur de journaux supprimait le dossier `user` vide sous `/var/log/hive`.
- Résolution d'un problème affectant les importations Hive qui avait une incidence sur le partitionnement et entraînait une erreur lors de l'importation.

- Résolution d'un problème où EMRFS et s3-dist-cp ne géraient pas correctement les noms de compartiment qui contenaient des points.
- Modification d'un comportement dans EMRFS de sorte que dans les compartiments activés pour la gestion des versions, le fichier marqueur `_$folder$` ne soit pas créé en permanence, ce qui peut contribuer à l'amélioration des performances de tels compartiments.
- Modification d'un comportement dans EMRFS pour ne pas utiliser pas des fichiers d'instruction à l'exception des cas où le chiffrement côté client est activé. Si vous souhaitez supprimer des fichiers d'instruction tout en utilisant le chiffrement côté client, vous pouvez définir la propriété d'emrfs-site.xml `fs.s3.cse.cryptoStorageMode.deleteInstructionFiles.enabled` sur true.
- Modification de l'agrégation de journaux YARN pour conserver les journaux dans la destination d'agrégation pendant deux jours. La destination par défaut est le stockage HDFS de votre cluster. Si vous souhaitez modifier cette durée, remplacez la valeur `yarn.log-aggregation.retain-seconds` à l'aide de la classification de configuration `yarn-site` lorsque vous créez votre cluster. Comme toujours, vous pouvez enregistrer vos journaux d'applications dans Amazon S3 à l'aide du paramètre `log-uri` lorsque vous créez votre cluster.

Correctifs appliqués

Les correctifs suivants de projets open source ont été inclus dans cette version :

- HIVE-9655
- HIVE-9183
- HADOOP-12810

Version 4.3.0

Date de parution : 19 janvier 2016

Fonctionnalités

Les fonctions suivantes sont disponibles dans cette version :

- Mise à niveau vers Hadoop 2.7.1
- Mise à niveau vers Spark 1.6.0
- Mise à niveau de Ganglia vers la version 3.7.2
- Mise à niveau vers de Presto vers la version 0.130

Amazon EMR a apporté quelques modifications au paramètre `spark.dynamicAllocation.enabled` lorsque celui-ci est défini sur true ; la valeur par défaut est false. Lorsque ce paramètre est défini sur true, cela a un effet sur les valeurs par défaut définies par le paramètre `maximizeResourceAllocation`.

- Si `spark.dynamicAllocation.enabled` est défini sur true, `spark.executor.instances` n'est pas défini par `maximizeResourceAllocation`.
- Le paramètre `spark.driver.memory` est désormais configuré en fonction des types d'instances du cluster d'une manière similaire à la façon dont `spark.executors.memory` est défini. Cependant, étant donné que l'application de pilote Spark peut s'exécuter soit sur le maître, soit sur une des instances principales (par exemple, dans le client YARN et les modes cluster, respectivement), le paramètre `spark.driver.memory` est défini en fonction du type d'instance du plus petit des types d'instances entre ces deux groupes d'instances.
- Le paramètre `spark.default.parallelism` est désormais défini sur deux fois le nombre de cœurs de processeurs disponibles pour les conteneurs YARN. Dans les versions précédentes, c'était la moitié de cette valeur.
- Les calculs de la surcharge de mémoire réservée pour les processus Spark YARN ont été ajustés pour être plus précis, ce qui se traduit par une petite augmentation de la quantité de mémoire disponible pour Spark (c'est-à-dire, `spark.executor.memory`).

Problèmes connus résolus depuis les versions précédentes

- L'agrégation de journaux YARN est maintenant activée par défaut.
- Résolution d'un problème où les journaux ne pouvaient pas être transmis vers le compartiment de journaux d'un cluster Amazon S3 lorsque l'agrégation de journaux YARN était activée.
- Les tailles de conteneur YARN ont maintenant un nouveau minimum de 32 sur tous les types de nœuds.
- Correction d'un problème lié à Ganglia qui provoquait un excès de disque I/O sur le nœud principal dans les grands clusters.
- Résolution d'un problème qui empêchait les journaux des applications d'être transmis à Amazon S3 lorsqu'un cluster s'arrêtait.
- Résolution d'un problème dans l'interface de ligne de commande EMRFS qui entraînait l'échec de certaines commandes.
- Correction d'un problème lié à Zeppelin qui empêchait le chargement des dépendances dans le sous-jacent. SparkContext

- Résolution d'un problème provoqué par une tentative de redimensionnement pour ajouter des instances.
- Résolution d'un problème dans Hive où CREATE TABLE AS SELECT effectuait des appels de liste excessifs vers Amazon S3.
- Résolution d'un problème où les grands clusters n'étaient pas provisionnés correctement lorsque Hue, Oozie et Ganglia étaient installés.
- Résolution d'un problème dans s3-dist-cp où un code de sortie zéro était renvoyé même en cas d'échec avec une erreur.

Correctifs appliqués

Les correctifs suivants de projets open source ont été inclus dans cette version :

- OOZIE-2402
- HIVE-12502
- HIVE-10631
- HIVE-12213
- HIVE-10559
- HIVE-12715
- HIVE-10685

Version 4.2.0

Date de parution : 18 novembre 2015

Fonctionnalités

Les fonctions suivantes sont disponibles dans cette version :

- Ajout de la prise en charge de Ganglia
- Mise à niveau vers Spark 1.5.2
- Mise à niveau vers Presto 0.125
- Mise à niveau d'Oozie vers la version 4.2.0
- Mise à niveau de Zeppelin vers la version 0.5.5
- Mise à niveau de la AWS SDK pour Java version 1.10.27

Problèmes connus résolus depuis les versions précédentes

- Résolution d'un problème où l'interface de ligne de commande EMRFS n'utilisait pas le nom de table de métadonnées par défaut.
- Résolution d'un problème rencontré lors de l'utilisation de tables basée sur ORC dans Amazon S3.
- Résolution d'un problème rencontré avec une incompatibilité de version Python dans la configuration de Spark.
- Résolution d'un problème où un état de nœud YARN n'était pas signalé en raison de problèmes de DNS pour des clusters dans un VPC.
- Résolution d'un problème rencontré lorsque YARN mettait hors service des nœuds, qui se traduisait le blocage d'applications ou l'incapacité de planifier de nouvelles applications.
- Résolution d'un problème rencontré lorsque des clusters prenaient fin avec l'état TIMED_OUT_STARTING.
- Résolution d'un problème rencontré lors de l'inclusion de la dépendance EMRFS Scala dans d'autres versions. La dépendance Scala a été supprimée.

Configuration des applications

Pour remplacer les configurations par défaut d'une application, vous pouvez fournir un objet de configuration. Vous pouvez soit utiliser une syntaxe abrégée pour fournir la configuration, soit faire référence à l'objet de configuration dans un fichier JSON. Les objets de configuration sont composés d'une classification, de propriétés et de configurations imbriquées en option. Les propriétés correspondent aux paramètres de l'application que vous souhaitez modifier. Vous pouvez spécifier plusieurs classifications pour plusieurs applications d'un seul objet JSON.

> **Warning**
>
> Les opérations de l'API Amazon EMR Describe et List émettent des paramètres personnalisés et configurables, qui sont utilisés dans le cadre des flux de travail Amazon EMR, en texte clair. Pour fournir des informations sensibles, telles que des mots de passe, dans ces paramètres, consultez Stocker les données de configuration sensibles dans AWS Secrets Manager.

Les classifications de configuration qui sont disponibles varient d'une version d'Amazon EMR à l'autre. Pour obtenir la liste des classifications de configuration prises en charge dans une version de version particulière, reportez-vous à la page de cette version sous À propos des versions Amazon EMR.

Voici un exemple de fichier JSON pour une liste de configurations.

```
[
  {
    "Classification": "core-site",
    "Properties": {
      "hadoop.security.groups.cache.secs": "250"
    }
  },
  {
    "Classification": "mapred-site",
    "Properties": {
      "mapred.tasktracker.map.tasks.maximum": "2",
      "mapreduce.map.sort.spill.percent": "0.90",
      "mapreduce.tasktracker.reduce.tasks.maximum": "5"
    }
```

```
  }
]
```

Une classification de configuration correspond généralement à un fichier de configuration spécifique à l'application. Par exemple, la classification `hive-site` est mappée aux paramètres du fichier de configuration `hive-site.xml` pour Hive. Une exception à cette règle est l'action d'amorçage `configure-daemons`, qui n'est plus prise en charge et qui est utilisée pour définir des paramètres d'environnement tels que `--namenode-heap-size`. Les options de ce type sont intégrées dans les classifications `hadoop-env` et `yarn-env` avec leurs propres classifications d'exportation imbriquées. Si des classifications se terminent par `env`, utilisez la sous-classification d'exportation.

`s3get` est une autre exception qui est utilisée pour placer un objet `EncryptionMaterialsProvider` client sur chaque nœud d'un cluster pour l'utiliser dans le chiffrement côté client. Une option a été ajoutée à la classification `emrfs-site` à cet effet.

Voici un exemple de la classification `hadoop-env`.

```
[
  {
    "Classification": "hadoop-env",
    "Properties": {

    },
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "HADOOP_DATANODE_HEAPSIZE": "2048",
          "HADOOP_NAMENODE_OPTS": "-XX:GCTimeRatio=19"
        },
        "Configurations": [

        ]
      }
    ]
  }
]
```

Voici un exemple de la classification yarn-env.

```
[
```

```
  {
    "Classification": "yarn-env",
    "Properties": {

    },
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "YARN_RESOURCEMANAGER_OPTS": "-Xdebug -Xrunjdwp:transport=dt_socket"
        },
        "Configurations": [

        ]
      }
    ]
  }
]
```

Les paramètres suivants n'appartiennent pas à un fichier de configuration, mais sont utilisés par Amazon EMR pour configurer potentiellement plusieurs paramètres en votre nom.

Paramètres définis par Amazon EMR

Application	Classification d'étiquette de version	Propriétés valides	Utilisation
Spark	`spark`	`maximizeResourceAllocation`	Configurez des programmes d'exécution pour utiliser le maximum de ressources de chaque nœud.

Rubriques

- Configuration des applications lorsque vous créez un cluster
- Reconfigurer un groupe d'instances dans un cluster en cours d'exécution
- Stocker les données de configuration sensibles dans AWS Secrets Manager
- Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique

Configuration des applications lorsque vous créez un cluster

Lorsque vous créez un cluster, vous pouvez remplacer les configurations par défaut des applications à l'aide de la console Amazon EMR,AWS CLI() ou AWS Command Line Interface AWS le kit SDK.

Pour remplacer la configuration par défaut d'une application, vous devez spécifier des valeurs personnalisées dans une classification de configuration. Une classification de configuration correspond à un fichier XML de configuration pour une application, tel que `hive-site.xml`.

Les classifications de configuration varient selon la version d'Amazon EMR. Pour obtenir une liste des classifications de configuration disponibles dans une version spécifique, consultez la page des détails de la version. Par exemple, Amazon EMR version 6.4.0.

Fournir une configuration dans la console lorsque vous créez un cluster

Pour fournir une configuration, accédez à la page Créer un cluster et développez les paramètres du logiciel. Vous pouvez ensuite saisir la configuration directement à l'aide de JSON ou d'une syntaxe raccourcie présentée dans le texte ombré dans la console. Sinon, vous pouvez fournir un URI Amazon S3 pour un fichier avec un objet JSON `Configurations`.

Pour fournir une configuration pour un groupe d'instances, choisissez un cluster dans votre liste de clusters, puis sélectionnez l'onglet Configurations. Dans le tableau des configurations des groupes d'instances, choisissez le groupe d'instances à modifier, puis choisissez Reconfigurer.

Fournir une configuration à l'aide de l' AWS CLI option lorsque vous créez un cluster

Vous pouvez fournir une configuration à create-cluster en fournissant un chemin d'accès à un fichier JSON stocké localement ou dans Amazon S3. L'exemple suivant suppose que vous utilisiez les rôles par défaut pour Amazon EMR et que ces rôles ont été créés. Si vous devez créer les rôles, exécutez d'abord `aws emr create-default-roles`.

Si votre configuration se trouve dans votre répertoire local, vous pouvez utiliser l'exemple de commande suivant.

```
aws emr create-cluster --use-default-roles --release-label emr-7.9.0 --applications
 Name=Hive \
--instance-type m5.xlarge --instance-count 3 --configurations file://./
configurations.json
```

Si votre configuration se trouve dans un chemin Amazon S3, vous devez configurer la solution de contournement suivante avant de transmettre le chemin Amazon S3 à la commande `create-cluster`.

```
#!/bin/sh
# Assume the ConfigurationS3Path is not public, and its present in the same AWS account
 as the EMR cluster
ConfigurationS3Path="s3://amzn-s3-demo-bucket/config.json"
# Get a presigned HTTP URL for the s3Path
ConfigurationURL=`aws s3 presign $ConfigurationS3Path --expires-in 300`
# Fetch the presigned URL, and minify the JSON so that it spans only a single line
Configurations=`curl $ConfigurationURL | jq -c .`
aws emr create-cluster --use-default-roles --release-label emr-5.34.0 --instance-type
 m5.xlarge --instance-count 2 --applications Name=Hadoop Name=Spark --configurations
 $Configurations
```

Fournir une configuration à l'aide du kit SDK Java lors de la création d'un cluster

L'extrait de programme suivant montre comment fournir une configuration à l'aide de AWS SDK pour Java.

```
Application hive = new Application().withName("Hive");

Map<String,String> hiveProperties = new HashMap<String,String>();
 hiveProperties.put("hive.join.emit.interval","1000");
 hiveProperties.put("hive.merge.mapfiles","true");

Configuration myHiveConfig = new Configuration()
 .withClassification("hive-site")
 .withProperties(hiveProperties);

RunJobFlowRequest request = new RunJobFlowRequest()
 .withName("Create cluster with ReleaseLabel")
 .withReleaseLabel("emr-5.20.0")
 .withApplications(hive)
 .withConfigurations(myHiveConfig)
 .withServiceRole("EMR_DefaultRole")
 .withJobFlowRole("EMR_EC2_DefaultRole")
 .withInstances(new JobFlowInstancesConfig()
  .withEc2KeyName("myEc2Key")
```

```
  .withInstanceCount(3)
  .withKeepJobFlowAliveWhenNoSteps(true)
  .withMasterInstanceType("m4.large")
  .withSlaveInstanceType("m4.large")
);
```

Reconfigurer un groupe d'instances dans un cluster en cours d'exécution

Avec Amazon EMR version 5.21.0 et ultérieure, vous pouvez reconfigurer les applications de cluster et spécifier des classifications de configuration supplémentaires pour chaque groupe d'instances dans un cluster en cours d'exécution. Pour ce faire, vous pouvez utiliser la console Amazon EMR, l' AWS Command Line Interface (AWS CLI) ou le AWS kit SDK.

Lorsque vous mettez à jour la configuration d'une application pour un groupe d'instances dans la nouvelle console Amazon EMR, la console tente de fusionner la nouvelle configuration avec la configuration existante pour créer une nouvelle configuration active. Dans le cas inhabituel où Amazon EMR ne parvient pas à fusionner la configuration, la console vous avertit.

Après avoir soumis une demande de reconfiguration pour un groupe d'instances, Amazon EMR attribue un numéro de version à la nouvelle spécification de configuration. Vous pouvez suivre le numéro de version d'une configuration, ou l'état d'un groupe d'instances, en consultant les CloudWatch événements. Pour plus d'informations, consultez la section Surveiller CloudWatch les événements.

> **Note**
>
> Vous pouvez uniquement remplacer, et non supprimer, les configurations de cluster spécifiées lors de la création du cluster. S'il existe des différences entre la configuration existante et le fichier que vous fournissez, Amazon EMR rétablit les configurations modifiées manuellement, telles que les configurations que vous avez modifiées lors de la connexion à votre cluster via SSH, aux valeurs par défaut du cluster pour le groupe d'instances spécifié.

Considérations à prendre en compte lors de la reconfiguration d'un groupe d'instances

Actions de reconfiguration

Lorsque vous soumettez une demande de reconfiguration à l'aide de la console Amazon EMR,AWS CLI() ou AWS Command Line Interface du kit SDK, Amazon EMR vérifie AWS le fichier de configuration existant sur le cluster. S'il existe des différences entre la configuration existante et le fichier que vous fournissez, Amazon EMR lance des actions de reconfiguration, redémarre certaines applications et réinitialise toutes les configurations modifiées manuellement, telles que les configurations que vous avez modifiées lorsque vous êtes connecté à votre cluster via SSH, aux valeurs par défaut du cluster pour le groupe d'instances spécifié.

> **Note**
>
> Amazon EMR exécute certaines actions par défaut lors de chaque reconfiguration de groupe d'instances. Ces actions par défaut peuvent entrer en conflit avec les personnalisations de cluster que vous avez effectuées et entraîner des échecs de reconfiguration. Pour plus d'informations sur le dépannage des échecs de reconfiguration, consultez Résoudre les problèmes de reconfiguration des groupes d'instances.

Amazon EMR lance également des actions de reconfiguration pour les classifications de configuration que vous spécifiez dans votre demande. Pour obtenir la liste complète de ces actions, consultez la section Classifications de configuration correspondant à la version d'Amazon EMR que vous utilisez. Par exemple, Classifications de configuration 6.2.0.

> **Note**
>
> Le guide de mise à jour d'Amazon EMR répertorie uniquement les actions de reconfiguration à partir des versions 5.32.0 et 6.2.0 d'Amazon EMR.

Interruption du service

Amazon EMR suit un processus continu pour reconfigurer les instances dans les groupes d'instances Task et Core. Seules 10 % des instances d’un groupe d'instances peuvent être modifiées et redémarrées à un moment donné. Ce processus est plus long à s’exécuter, mais il

permet de réduire le risque de défaillance potentielle des applications dans un cluster en cours d'exécution.

Pour exécuter des tâches YARN lors d'un redémarrage de YARN, vous pouvez soit créer un cluster Amazon EMR avec plusieurs nœuds primaires, soit définir `yarn.resourcemanager.recovery.enabled` sur `true` dans votre classification de configuration `yarn-site`. Pour plus d'informations sur l'utilisation de plusieurs nœuds principaux, consultez YARN à haute disponibilité ResourceManager.

Validation des applications

Amazon EMR vérifie que chaque application du cluster est en cours d'exécution après le redémarrage de la reconfiguration. Si une application n'est pas disponible, l'opération de reconfiguration globale échoue. Si une opération de reconfiguration échoue, Amazon EMR inverse les paramètres de configuration pour revenir à la version de travail précédente.

> **Note**
>
> Pour éviter tout échec de reconfiguration, nous vous recommandons de n'installer sur votre cluster que les applications que vous comptez utiliser. Nous vous recommandons également de vous assurer que toutes les applications de cluster fonctionnent correctement avant de soumettre une demande de reconfiguration.

Types de reconfiguration

Vous pouvez reconfigurer un groupe d'instances de deux manières suivantes :

- Overwrite. Méthode de reconfiguration par défaut et la seule disponible dans les versions Amazon EMR antérieures aux versions 5.35.0 et 6.6.0. Cette méthode de reconfiguration remplace sans distinction tous les fichiers du cluster par le nouvel ensemble de configuration soumis. La méthode efface toutes les modifications apportées aux fichiers de configuration en dehors de l'API de reconfiguration.
- Fusionner. Méthode de reconfiguration prise en charge pour les versions 5.35.0 et 6.6.0 d'Amazon EMR et versions ultérieures, à l'exception de la console Amazon EMR, où aucune version ne la prend en charge. Cette méthode de reconfiguration fusionne les configurations récemment soumises avec les configurations qui existent déjà sur le cluster. Cette option ajoute ou modifie uniquement les nouvelles configurations que vous soumettez. Il préserve les configurations existantes.

> **Note**
>
> Amazon EMR continue de remplacer certaines configurations Hadoop essentielles dont il a besoin pour garantir le bon fonctionnement du service.

Limites

Lorsque vous reconfigurez un groupe d'instances dans un cluster en cours d'exécution, tenez compte des limitations suivantes :

- Les applications autres que Yarn peuvent échouer au redémarrage ou provoquer des problèmes de cluster, en particulier si les applications ne sont pas correctement configurées. Les clusters approchant le maximum d'utilisation de la mémoire et du processeur peuvent rencontrer des problèmes après le redémarrage. Cela est particulièrement vrai pour le groupe d'instances principal.
- Vous ne pouvez pas envoyer de demande de reconfiguration lorsqu'un groupe d'instances est redimensionné. Si une reconfiguration est initiée pendant le redimensionnement d'un groupe d'instances, la reconfiguration ne peut pas démarrer tant que le redimensionnement du groupe d'instances n'est pas terminé, et vice versa.
- Après avoir reconfiguré un groupe d'instances, Amazon EMR redémarre les applications afin de permettre la prise en compte des nouvelles configurations. Des échecs de tâches ou d'autres comportements inattendus des applications peuvent se produire si les applications sont utilisées durant la reconfiguration.
- Si une reconfiguration échoue pour un groupe d'instances, Amazon EMR réinitialise les paramètres de configuration vers la précédente version fonctionnelle. Si le processus de réinitialisation échoue, vous devez soumettre une nouvelle demande `ModifyInstanceGroup` pour récupérer le groupe d'instances à partir de l'état `SUSPENDED`.
- Les demandes de reconfiguration pour les classifications de configuration Phoenix ne sont prises en charge que dans Amazon EMR version 5.23.0 et ultérieure, et ne sont pas prises en charge dans Amazon EMR version 5.21.0 ou 5.22.0.
- Les demandes de reconfiguration pour les classifications de HBase configuration ne sont prises en charge que dans les versions 5.30.0 et ultérieures d'Amazon EMR, et non dans les versions 5.23.0 à 5.29.0 d'Amazon EMR.

- Amazon EMR prend en charge les demandes de reconfiguration d'applications sur un cluster Amazon EMR comportant plusieurs nœuds primaires uniquement dans les versions 5.27.0 et ultérieures d'Amazon EMR.
- La reconfiguration de la classification `hdfs-encryption-zones` ou de l'une des classifications de configuration Hadoop KMS n'est pas prise en charge sur un cluster Amazon EMR comportant plusieurs nœuds primaires.
- Amazon EMR ne prend actuellement pas en charge certaines demandes de reconfiguration de l'ordonnanceur de capacité qui nécessitent le redémarrage de YARN. ResourceManager Par exemple, vous ne pouvez pas supprimer complètement une file d'attente.

Reconfigurer un groupe d'instances dans la console

> **Note**
>
> La console Amazon EMR ne prend pas en charge les reconfigurations de type Merge.

1. Ouvrez la console Amazon EMR à l'adresse /emr https://console.aws.amazon.com
2. Dans la liste des clusters sous Nom, sélectionnez le cluster actif que vous souhaitez reconfigurer.
3. Ouvrez la page des détails du cluster et accédez à l'onglet Configurations.
4. Dans la liste déroulante Filter (Filtre), sélectionnez le groupe d'instances que vous souhaitez reconfigurer.
5. Dans le menu déroulant Reconfigurer, choisissez Modifier dans la table ou Modifier dans le fichier JSON.

   - Modifier dans la table – Dans la table de classification de configuration, modifiez la propriété et la valeur des configurations existantes ou choisissez Ajouter une configuration pour fournir des classifications de configuration supplémentaires.
   - Modification dans un fichier JSON – Saisissez la configuration directement dans JSON, ou utilisez une syntaxe sténographique (démontrée dans le texte ombré). Sinon, fournissez un URI Amazon S3 pour un fichier avec un objet JSON `Configurations`.

> **Note**
>
> La colonne Source de la table des classifications de configuration indique si la configuration est fournie lorsque vous créez un cluster ou lorsque vous spécifiez des configurations supplémentaires pour ce groupe d'instances. Vous pouvez modifier les configurations d'un groupe d'instances à partir de ces deux sources. Vous ne pouvez pas supprimer les configurations de cluster initiales, mais vous pouvez les remplacer pour un groupe d'instances.
>
> Vous pouvez également ajouter ou modifier les classifications de configuration imbriquées directement dans la table. Par exemple, pour fournir une sous-classification `export` supplémentaire de `hadoop-env`, ajoutez une classification de configuration `hadoop.export` dans la table. Ensuite, indiquez une propriété et une valeur spécifiques pour cette classification.

6. (Facultatif) Sélectionnez Apply this configuration to all active instance groups (Appliquer cette configuration à tous les groupes d'instances actifs).
7. Enregistrez les Modifications.

Reconfigurer un groupe d'instances à l'aide de la CLI

Utilisez la commande modify-instance-groups pour spécifier une nouvelle configuration pour un groupe d'instances dans un cluster en cours d'exécution.

> **Note**
>
> Dans les exemples suivants, remplacez *<j-2AL4XXXXXX5T9>* par votre ID de cluster et *<ig-1xxxxxxx9>* par votre ID de groupe d'instances.

Example – Remplacer une configuration pour un groupe d'instances

L'exemple suivant fait référence à un fichier JSON de configuration appelé `instanceGroups.json` pour modifier la propriété du vérificateur de l'état des NodeManager disques de YARN pour un groupe d'instances.

1. Préparez votre classification de configuration et enregistrez-la comme `instanceGroups.json` dans le répertoire où vous allez exécuter la commande.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.disk-health-checker.enable":"true",
               "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. Exécutez la commande suivante.

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
--instance-groups file://instanceGroups.json
```

Example – Ajoute une configuration à un groupe d'instances

Si vous voulez ajouter une configuration à un groupe d'instances, vous devez inclure toutes les configurations précédemment spécifiées pour ce groupe d'instances dans votre nouvelle demande `ModifyInstanceGroup`. Dans le cas contraire, les configurations précédemment spécifiées sont supprimées.

L'exemple suivant ajoute une propriété pour le vérificateur de mémoire NodeManager virtuelle de YARN. La configuration inclut également les valeurs précédemment spécifiées pour le vérificateur de santé du NodeManager disque YARN afin que les valeurs ne soient pas remplacées.

1. Préparez le contenu suivant dans `instanceGroups.json` et enregistrez-le dans le même répertoire que celui où vous exécuterez la commande.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.disk-health-checker.enable":"true",
               "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0",
               "yarn.nodemanager.vmem-check-enabled":"true",
               "yarn.nodemanager.vmem-pmem-ratio":"3.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. Exécutez la commande suivante.

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
--instance-groups file://instanceGroups.json
```

Example – Ajoutez une configuration à un groupe d'instances avec une reconfiguration de type Merge

Lorsque vous souhaitez utiliser la méthode de reconfiguration Overwrite par défaut pour ajouter une configuration, vous devez inclure toutes les configurations précédemment spécifiées pour ce groupe d'instances dans votre nouvelle demande `ModifyInstanceGroup`. Dans le cas contraire, Overwrite supprime les configurations que vous avez spécifiées précédemment. Vous n'avez pas besoin de le faire avec la reconfiguration de Merge. Vous devez plutôt vous assurer que votre demande inclut uniquement les nouvelles configurations.

L'exemple suivant ajoute une propriété pour le vérificateur de mémoire NodeManager virtuelle de YARN. Comme il s'agit d'une reconfiguration de type Merge, elle ne remplace pas les valeurs précédemment spécifiées pour le vérificateur de santé du NodeManager disque YARN.

1. Préparez le contenu suivant dans `instanceGroups.json` et enregistrez-le dans le même répertoire que celui où vous exécuterez la commande.

```
[
   {"InstanceGroupId":"<ig-1xxxxxxx9>",
    "ReconfigurationType" :"MERGE",
      "Configurations":[
         {"Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.vmem-check-enabled":"true",
               "yarn.nodemanager.vmem-pmem-ratio":"3.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. Exécutez la commande suivante.

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
--instance-groups file://instanceGroups.json
```

Example – Supprimer une configuration pour un groupe d'instances

Pour supprimer une configuration pour un groupe d'instances, soumettez une nouvelle demande de reconfiguration excluant la configuration précédente.

> **Note**
>
> Vous pouvez uniquement remplacer la configuration de cluster initiale. Vous ne pouvez pas la supprimer.

Par exemple, pour supprimer la configuration du vérificateur d'état des NodeManager disques de YARN de l'exemple précédent, soumettez une nouvelle demande `instanceGroups.json` avec le contenu suivant.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.vmem-check-enabled":"true",
               "yarn.nodemanager.vmem-pmem-ratio":"3.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

Note

Pour supprimer toutes les configurations de votre dernière demande de reconfiguration, soumettez une demande de reconfiguration avec un ensemble de configurations vide. Par exemple,

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[]
   }
]
```

Example – Reconfigurez et redimensionnez un groupe d'instances en une seule demande

L'exemple JSON suivant montre comment reconfigurer et redimensionner un groupe d'instances dans la même demande.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
```

```
      "InstanceCount":5,
      "EC2InstanceIdsToTerminate":["i-123"],
      "ForceShutdown":true,
      "ShrinkPolicy":{
         "DecommissionTimeout":10,
         "InstanceResizePolicy":{
            "InstancesToTerminate":["i-123"],
            "InstancesToProtect":["i-345"],
            "InstanceTerminationTimeout":20
         }
      },
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Configurations":[],
            "Properties":{
               "yarn.nodemanager.disk-health-checker.enable":"true",
               "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0"
            }
         }
      ]
   }
]
```

Reconfigurer un groupe d'instances à l'aide du SDK Java

> **Note**
>
> Dans les exemples suivants, remplacez *<j-2AL4XXXXXX5T9>* par votre ID de cluster et *<ig-1xxxxxxx9>* par votre ID de groupe d'instances.

L'extrait de code suivant fournit une nouvelle configuration pour un groupe d'instances utilisant AWS SDK pour Java.

```
AWSCredentials credentials = new BasicAWSCredentials("access-key", "secret-key");
AmazonElasticMapReduce emr = new AmazonElasticMapReduceClient(credentials);

Map<String,String> hiveProperties = new HashMap<String,String>();
```

```
hiveProperties.put("hive.join.emit.interval","1000");
hiveProperties.put("hive.merge.mapfiles","true");

Configuration configuration = new Configuration()
    .withClassification("hive-site")
    .withProperties(hiveProperties);

InstanceGroupModifyConfig igConfig = new InstanceGroupModifyConfig()
    .withInstanceGroupId("<ig-1xxxxxxx9>")
    .withReconfigurationType("MERGE");
    .withConfigurations(configuration);

ModifyInstanceGroupsRequest migRequest = new ModifyInstanceGroupsRequest()
    .withClusterId("<j-2AL4XXXXXX5T9>")
    .withInstanceGroups(igConfig);

emr.modifyInstanceGroups(migRequest);
```

L'extrait de code suivant supprime une configuration précédemment spécifiée pour un groupe d'instances en fournissant un tableau vide de configurations.

```
List<Configuration> configurations = new ArrayList<Configuration>();

InstanceGroupModifyConfig igConfig = new InstanceGroupModifyConfig()
    .withInstanceGroupId("<ig-1xxxxxxx9>")
    .withConfigurations(configurations);

ModifyInstanceGroupsRequest migRequest = new ModifyInstanceGroupsRequest()
    .withClusterId("<j-2AL4XXXXXX5T9>")
    .withInstanceGroups(igConfig);

emr.modifyInstanceGroups(migRequest);
```

Résoudre les problèmes de reconfiguration des groupes d'instances

Si le processus de reconfiguration d'un groupe d'instances échoue, Amazon EMR annule la reconfiguration et enregistre un message d'échec à l'aide d'un événement Amazon. CloudWatch L'événement fournit un bref résumé de l'échec de la reconfiguration. Il répertorie les instances pour lesquelles la reconfiguration a échoué et les messages d'échec correspondants. Voici un exemple de message d'échec.

```
The reconfiguration operation for instance group ig-1xxxxxxx9 in Amazon EMR
 cluster j-2AL4XXXXXX5T9 (ExampleClusterName)
failed at 2021-01-01 00:00 UTC and took 2 minutes to fail. Failed configuration version
 is example12345.
Failure message: Instance i-xxxxxxx1, i-xxxxxxx2, i-xxxxxxx3 failed with message "This
 is an example failure message".
```

Pour recueillir davantage de données sur un échec de reconfiguration, vous pouvez consulter les journaux de provisionnement des nœuds. Cela est particulièrement utile lorsque vous recevez un message comme celui-ci.

```
i-xxxxxxx1 failed with message “Unable to complete transaction and some changes were
 applied.”
```

On the node

Pour accéder aux journaux de provisionnement des nœuds en vous connectant à un nœud

1. Utilisez SSH pour vous connecter au nœud sur lequel la reconfiguration a échoué. Pour obtenir des instructions, consultez Connect to your Linux instance dans le guide de EC2 l'utilisateur Amazon pour les instances Linux.
2. Accédez au répertoire suivant, qui contient les fichiers journaux de provisionnement des nœuds.

   ```
   /mnt/var/log/provision-node/
   ```

3. Ouvrez le sous-répertoire `reports` et recherchez le rapport de provisionnement des nœuds pour votre reconfiguration. Le `reports` répertoire organise les journaux par numéro de version de reconfiguration, identifiant unique universel (UUID), adresse IP de l' EC2 instance Amazon et horodatage. Chaque rapport est un fichier YAML compressé qui contient des informations détaillées sur le processus de reconfiguration.

   Vous trouverez ci-dessous un exemple de nom de fichier et de chemin d'accès.

   ```
   /reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxx/ip-10-73-xxx-
   xxx.ec2.internal/202104061715.yaml.gz
   ```

4. Vous pouvez examiner un rapport à l'aide d'un visualiseur de fichiers tel que `zless`, comme dans l'exemple suivant.

```
zless 202104061715.yaml.gz
```

Amazon S3

Pour accéder aux journaux de provisionnement des nœuds à l'aide d'Amazon S3

1. Connectez-vous à la AWS Management Console et ouvrez la console Amazon S3 à l'adresse https://console.aws.amazon.com/s3/.
2. Ouvrez le compartiment Amazon S3 que vous avez spécifié lorsque vous avez configuré le cluster pour archiver les fichiers journaux.
3. Accédez au dossier suivant, qui contient les fichiers journaux de provisionnement des nœuds :

   ```
   amzn-s3-demo-bucket/elasticmapreduce/<cluster id>/node/<instance id>/provision-node/
   ```

4. Ouvrez le dossier `reports` et recherchez le rapport de provisionnement des nœuds pour votre reconfiguration. Le `reports` dossier organise les journaux par numéro de version de reconfiguration, identifiant unique universel (UUID), adresse IP de l' EC2 instance Amazon et horodatage. Chaque rapport est un fichier YAML compressé qui contient des informations détaillées sur le processus de reconfiguration.

   Vous trouverez ci-dessous un exemple de nom de fichier et de chemin d'accès.

   ```
   /reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxx/ip-10-73-xxx-xxx.ec2.internal/202104061715.yaml.gz
   ```

5. Pour consulter un fichier journal, vous pouvez le télécharger depuis Amazon S3 sur votre machine locale sous forme de fichier texte. Pour obtenir des instructions, consultez la section Téléchargement d'un objet.

Chaque fichier journal contient un rapport de provisionnement détaillé pour la reconfiguration associée. Pour trouver les informations relatives aux messages d'erreur, vous pouvez rechercher le niveau de journalisation `err` d'un rapport. Le format du rapport dépend de la version d'Amazon EMR de votre cluster.

L'exemple suivant montre les informations d'erreur pour les versions d'Amazon EMR antérieures aux versions 5.32.0 et 6.2.0.

```
- !ruby/object:Puppet::Util::Log
      level: !ruby/sym err
      tags:
        - err
      message: "Example detailed error message."
      source: Puppet
      time: 2021-01-01 00:00:00.000000 +00:00
```

Les versions 5.32.0 et 6.2.0 et ultérieures d'Amazon EMR utilisent plutôt le format suivant.

```
- level: err
  message: 'Example detailed error message.'
  source: Puppet
  tags:
  - err
  time: '2021-01-01 00:00:00.000000 +00:00'
  file:
  line:
```

Stocker les données de configuration sensibles dans AWS Secrets Manager

Amazon EMR décrit et répertorie les opérations d'API qui émettent des données de configuration personnalisées (telles que `DescribeCluster` et `ListInstanceGroups`) en texte clair. Amazon EMR s'intègre à AWS Secrets Manager afin que vous puissiez stocker vos données dans Secrets Manager et utiliser l'ARN secret dans vos configurations. De cette manière, vous ne transmettez pas de données de configuration sensibles à Amazon EMR en clair et ne les exposez pas à des données de configuration sensibles à Amazon EMR en clair et ne les exposez pas à des données externes. APIs Si vous indiquez qu'une paire clé-valeur contient un ARN pour un secret stocké dans Secrets Manager, Amazon EMR récupère ce secret lorsqu'il envoie les données de configuration au cluster. Amazon EMR n'envoie pas l'annotation lorsqu'il utilise une option externe APIs pour afficher la configuration.

Créer un secret

Pour créer un secret, suivez les étapes de la section Création d'un secret AWS Secrets Manager dans le Guide de l'utilisateur AWS Secrets Manager . Dans l'étape 3, vous devez choisir le champ de texte en clair pour saisir votre valeur sensible.

Notez que si Secrets Manager autorise un secret à contenir jusqu'à 65 536 octets, Amazon EMR limite la longueur combinée de la clé de propriété (à l'exception de l'annotation) et de la valeur secrète récupérée à 1024 caractères.

Accordez à Amazon EMR l'accès pour récupérer le secret

Amazon EMR utilise un rôle de service IAM pour provisionner et gérer des clusters à votre place. Le rôle de service pour Amazon EMR définit les actions autorisées pour Amazon EMR lorsqu'il alloue des ressources et exécute des tâches de niveau service qui ne sont pas effectuées dans le contexte d'une instance Amazon exécutée au sein d'un cluster. EC2 Pour plus d'informations sur les rôles de service, consultez Rôle de service pour Amazon EMR (rôle EMR) et Personnaliser les rôles IAM.

Pour permettre à Amazon EMR de récupérer la valeur secrète depuis Secrets Manager, ajoutez la déclaration de politique suivante à votre rôle Amazon EMR lorsque vous lancez votre cluster.

```
{
   "Sid":"AllowSecretsRetrieval",
   "Effect":"Allow",
   "Action":"secretsmanager:GetSecretValue",
   "Resource":[
      "arn:aws:secretsmanager:<region>:<aws-account-id>:secret:<secret-name>"
   ]
}
```

Si vous créez le secret avec une gérée par le client AWS KMS key, vous devez également ajouter une `kms:Decrypt` autorisation au rôle Amazon EMR pour la clé que vous utilisez. Pour plus d'informations, consultez Authentification et contrôle d'accès pour AWS Secrets Manager dans le Guide de l'utilisateur AWS Secrets Manager .

Utiliser le secret dans une classification de configuration

Vous pouvez ajouter l'annotation `EMR.secret@` à n'importe quelle propriété de configuration pour indiquer que sa paire clé-valeur contient un ARN pour un secret stocké dans Secrets Manager.

L'exemple suivant montre comment fournir un ARN secret dans une classification de configuration :

```
{
   "Classification":"core-site",
   "Properties":{
      "presto.s3.access-key":"<sensitive-access-key>",
      "EMR.secret@presto.s3.secret-key":"arn:aws:secretsmanager:<region>:<aws-account-
id>:secret:<secret-name>"
   }
}
```

Lorsque vous créez votre cluster et soumettez votre configuration annotée, Amazon EMR valide les propriétés de configuration. Si votre configuration est valide, Amazon EMR supprime l'annotation de la configuration et récupère le secret auprès de Secrets Manager pour créer la configuration réelle avant de l'appliquer au cluster :

```
{
   "Classification":"core-site",
   "Properties":{
      "presto.s3.access-key":"<sensitive-access-key>",
      "presto.s3.secret-key":"<my-secret-key-retrieved-from-Secrets-Manager>"
   }
}
```

Lorsque vous appelez une action telle que `DescribeCluster`, Amazon EMR renvoie la configuration actuelle de l'application sur le cluster. Si une propriété de configuration d'application est marquée comme contenant un ARN secret, la configuration d'application renvoyée par l'appel `DescribeCluster` contient l'ARN et non la valeur secrète. Cela garantit que la valeur secrète n'est visible que sur le cluster :

```
{
   "Classification":"core-site",
   "Properties":{
      "presto.s3.access-key":"<sensitive-access-key>",
      "presto.s3.secret-key":"arn:aws:secretsmanager:<region>:<aws-account-
id>:secret:<secret-name>"
   }
}
```

Mettre à jour la valeur d'un secret

Amazon EMR récupère la valeur secrète d'une configuration annotée chaque fois que le groupe d'instances attaché démarre, se reconfigure ou se redimensionne. Vous pouvez utiliser Secrets Manager pour modifier la valeur d'un secret utilisé dans la configuration d'un cluster en cours d'exécution. Dans ce cas, vous pouvez envoyer une demande de reconfiguration à chaque groupe d'instances pour lequel vous souhaitez recevoir la valeur mise à jour. Pour plus d'informations sur la façon de reconfigurer un groupe d'instances et sur les éléments à prendre en compte lorsque vous le faites, consultez Reconfigurer un groupe d'instances dans un cluster en cours d'exécution.

Configuration des applications pour l'utilisation d'une machine virtuelle Java spécifique

Les versions d'Amazon EMR comportent différentes versions de machine virtuelle Java (JVM) par défaut. Cette page explique la prise en charge de la JVM pour différentes versions et applications.

Considérations

Pour plus d’informations sur les versions de Java prises en charge pour les applications, voir les pages relatives aux applications du Guide des versions d’Amazon EMR.

- Amazon EMR ne prend en charge que l’exécution d’une seule version d’exécution dans un cluster et ne prend pas en charge l’exécution de différents nœuds ou applications sur différentes versions d’exécution sur le même cluster.
- Pour les versions 7.x d’Amazon EMR, la machine virtuelle Java (JVM) par défaut est Java 17 pour les applications compatibles avec Java 17, à l’exception d’Apache Livy. Pour plus d’informations sur les versions de JDK prises en charge pour les applications, voir les pages dédiées du Guide des versions d’Amazon EMR.
- À partir d'Amazon EMR 7.1.0, Flink prend en charge Java 17 et est configuré sur Java 17 par défaut. Pour utiliser une version différente de l'environnement d'exécution Java, remplacez les paramètres définis dans`flink-conf`. Pour plus d'informations sur la configuration de Flink pour utiliser Java 8 ou Java 11, voir Configurer Flink pour fonctionner avec Java 11.
- Pour les versions 5.x et 6.x d’Amazon EMR, la machine virtuelle Java (JVM) par défaut est Java 8.
  - Pour les versions 6.12.0 et supérieures d'Amazon EMR, certaines applications prennent également en charge Java 11 et 17.

- Pour les versions 6.9.0 et supérieures d'Amazon EMR, Trino prend en charge Java 17 par défaut. Pour plus d'informations sur Java 17 avec Trino, consultez Trino updates to Java 17 sur le blog de Trino.

Tenez compte des considérations spécifiques à l'application suivantes lorsque vous choisissez votre version d'exécution :

Remarques de configuration Java spécifiques à l'application

Application	Notes sur la configuration de Java
Spark	Pour exécuter Spark avec une version de Java autre que celle par défaut, vous devez configurer à la fois Spark et Hadoop. Pour obtenir des exemples, consultez Remplacer la JVM. • Configurez `JAVA_HOME` dans `spark-env` pour mettre à jour le runtime Java des processus de l'instance principale. Par exemple, spark-submit, spark-shell et Spark History Server. • Modifiez la configuration Hadoop pour mettre à jour le runtime Java des exécuteurs Spark et de YARN `ApplicationMaster`
Spark RAPIDS	Vous pouvez exécuter RAPIDS avec la version Java configurée pour Spark.
Iceberg	Vous pouvez exécuter Iceberg avec la version Java configurée de l'application qui l'utilise.
Delta	Vous pouvez exécuter Delta avec la version Java configurée de l'application qui l'utilise.
Hudi	Vous pouvez exécuter Hudi avec la version Java configurée de l'application qui l'utilise.

Application	Notes sur la configuration de Java
Hadoop	Pour mettre à jour la JVM pour Flink, modifiez `hadoop-env` . Pour obtenir des exemples, consultez Remplacer la JVM.
Hive	Pour définir la version Java sur 11 ou 17 pour Hive, configurez le paramètre JVM Hadoop sur la version Java que vous souhaitez utiliser.
HBase	Pour mettre à jour la JVM pour HBase, modifiez`hbase-env` . Par défaut, Amazon EMR définit la HBase JVM en fonction de la configuration JVM pour Hadoop, sauf si vous remplacez les paramètres dans. `hbase-env` Pour obtenir des exemples, consultez Remplacer la JVM.
Flink	Pour mettre à jour la JVM pour Flink, modifiez `flink-conf` . Par défaut, Amazon EMR définit la JVM Flink en fonction de la configura tion JVM pour Hadoop, sauf si vous remplacez les paramètres dans `flink-conf` . Pour de plus amples informations, veuillez consulter Configurer Flink pour qu'il fonctionne avec Java 11.
Oozie	Pour configurer Oozie pour qu'il fonctionne sur Java 11 ou 17, configurez Oozie Server, le Oozie LauncherAM Launcher AM et modifiez les configurations de vos exécutables et tâches côté client. Vous pouvez également configure r `EmbeddedOozieServer` pour l'exécuter sur Java 17. Pour de plus amples informations, veuillez consulter Configurer la version Java pour Oozie.

Application	Notes sur la configuration de Java
Pig	Pig ne supporte que Java 8. Vous ne pouvez pas utiliser Java 11 ou 17 avec Hadoop et exécuter Pig sur le même cluster.

Remplacer la JVM

Pour remplacer le paramètre JVM d'une version d'Amazon EMR (par exemple, pour utiliser Java 17 avec un cluster utilisant Amazon EMR version 6.12.0), entrez le paramètre `JAVA_HOME` correspondant à sa classification d'environnement, qui est *`application`*`-env` pour toutes les applications sauf Flink. Pour Flink, la classification de l'environnement est `flink-conf`. Pour connaître les étapes de configuration de l'environnement d'exécution Java avec Flink, consultez Configurer Flink pour qu'il fonctionne avec Java 11.

Rubriques

- Remplacer le paramètre JVM par Apache Spark
- Remplacer le paramètre JVM par Apache HBase
- Remplacez le paramètre JVM par Apache Hadoop et Hive

Remplacer le paramètre JVM par Apache Spark

Lorsque vous utilisez Spark avec Amazon EMR versions 6.12 et supérieures, vous pouvez configurer l'environnement de telle sorte que les exécuteurs utilisent Java 11 ou 17. Et lorsque vous utilisez Spark avec des versions d'Amazon EMR inférieures à 5.x et que vous écrivez un pilote à soumettre en mode cluster, le pilote utilise Java 7. Vous pouvez toutefois configurer l'environnement de manière à ce que les exécuteurs utilisent Java 8.

Pour remplacer la JVM pour Spark, vous devez définir à la fois les classifications Hadoop et Spark.

```
{
"Classification": "hadoop-env",
        "Configurations": [
            {
"Classification": "export",
                "Configurations": [],
                "Properties": {
```

```
"JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
                }
            }
        ],
        "Properties": {}
    },
    {
"Classification": "spark-env",
        "Configurations": [
            {
"Classification": "export",
                "Configurations": [],
                "Properties": {
"JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
                }
            }
        ],
        "Properties": {}
    }
```

L'exemple suivant montre comment ajouter les paramètres de configuration requis pour EMR 7.0.0+ afin de garantir une utilisation cohérente de la version Java sur tous les composants.

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.executorEnv.JAVA_HOME": "/usr/lib/jvm/java-1.8.0",
      "spark.yarn.appMasterEnv.JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
    }
  },
  {
    "Classification": "hadoop-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
        }
      }
    ],
    "Properties": {}
```

```
  },
  {
    "Classification": "spark-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
        }
      }
    ],
    "Properties": {}
  }
]
```

Remplacer le paramètre JVM par Apache HBase

Pour configurer HBase pour utiliser Java 11, vous pouvez définir la configuration suivante lorsque vous lancez le cluster.

```
[
  {
    "Classification": "hbase-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/jre-11"
        }
      }
    ],
    "Properties": {}
  }
]
```

Remplacez le paramètre JVM par Apache Hadoop et Hive

L'exemple suivant montre comment configurer la JVM sur la version 17 pour Hadoop et Hive.

```
[
```

```
    {
        "Classification": "hadoop-env",
            "Configurations": [
                {
                    "Classification": "export",
                    "Configurations": [],
                    "Properties": {
                        "JAVA_HOME": "/usr/lib/jvm/jre-17"
                    }
                }
        ],
        "Properties": {}
    }
]
```

Ports de service

Les éléments suivants sont des ports de service YARN et HDFS. Ces paramètres reflètent les valeurs par défaut de Hadoop. D'autres services d'applications sont hébergés sur les ports par défaut, sauf indication contraire. Pour de plus amples informations, veuillez consulter la documentation du projet de l'application.

Paramètres de port pour YARN and HDFS

Paramètre	Nom d'hôte/Port
`fs.default.name`	par défaut (`hdfs://`*`emrDeterminedIP`* `:8020`)
`dfs.datanode.address`	par défaut (`0.0.0.0:50010`)
`dfs.datanode.http.address`	par défaut (`0.0.0.0:50075`)
`dfs.datanode.https.address`	par défaut (`0.0.0.0:50475`)
`dfs.datanode.ipc.address`	par défaut (`0.0.0.0:50020`)
`dfs.http.address`	par défaut (`0.0.0.0:50070`)
`dfs.https.address`	par défaut (`0.0.0.0:50470`)
`dfs.secondary.http.address`	par défaut (`0.0.0.0:50090`)

Paramètre	Nom d'hôte/Port
yarn.nodemanager.address	par défaut (${yarn.nodemanager.hostname}:0)
yarn.nodemanager.localizer.address	par défaut (${yarn.nodemanager.hostname}:8040)
yarn.nodemanager.webapp.address	par défaut (${yarn.nodemanager.hostname}:8042)
yarn.resourcemanager.address	par défaut (${yarn.resourcemanager.hostname}:8032)
yarn.resourcemanager.admin.address	par défaut (${yarn.resourcemanager.hostname}:8033)
yarn.resourcemanager.resource-tracker.address	par défaut (${yarn.resourcemanager.hostname}:8031)
yarn.resourcemanager.scheduler.address	par défaut (${yarn.resourcemanager.hostname}:8030)
yarn.resourcemanager.webapp.address	par défaut (${yarn.resourcemanager.hostname}:8088)
yarn.web-proxy.address	par défaut (no-value)
yarn.resourcemanager.hostname	*emrDeterminedIP*

> **Note**
>
> Le terme *emrDeterminedIP* est une adresse IP qui est générée par le plan de contrôle Amazon EMR. Dans la version la plus récente, cette convention a été supprimée, sauf pour les paramètres yarn.resourcemanager.hostname et fs.default.name.

Utilisateurs de l'application

Les applications exécutent des processus comme s'il s'agissait de leur propre utilisateur. Par exemple, Hive JVMs s'exécute en tant qu'utilisateur`hive`, MapReduce JVMs exécute en tant qu'utilisateur`mapred`, etc. Ceci est illustré dans l'exemple d'état des processus suivant.

```
USER       PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND
hive      6452  0.2  0.7 853684 218520 ?       Sl   16:32   0:13 /usr/lib/jvm/
java-openjdk/bin/java -Xmx256m -Dhive.log.dir=/var/log/hive -Dhive.log.file=hive-
metastore.log -Dhive.log.threshold=INFO -Dhadoop.log.dir=/usr/lib/hadoop
hive      6557  0.2  0.6 849508 202396 ?       Sl   16:32   0:09 /usr/lib/jvm/java-
openjdk/bin/java -Xmx256m -Dhive.log.dir=/var/log/hive -Dhive.log.file=hive-server2.log
 -Dhive.log.threshold=INFO -Dhadoop.log.dir=/usr/lib/hadoop/l
hbase     6716  0.1  1.0 1755516 336600 ?      Sl   Jun21   2:20 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_master -XX:OnOutOfMemoryError=kill -9 %p -Xmx1024m -ea -XX:
+UseConcMarkSweepGC -XX:+CMSIncrementalMode -Dhbase.log.dir=/var/
hbase     6871  0.0  0.7 1672196 237648 ?      Sl   Jun21   0:46 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_thrift -XX:OnOutOfMemoryError=kill -9 %p -Xmx1024m -ea -XX:
+UseConcMarkSweepGC -XX:+CMSIncrementalMode -Dhbase.log.dir=/var/
hdfs      7491  0.4  1.0 1719476 309820 ?      Sl   16:32   0:22 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_namenode -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-hdfs -
Dhadoop.log.file=hadoop-hdfs-namenode-ip-10-71-203-213.log -Dhadoo
yarn      8524  0.1  0.6 1626164 211300 ?      Sl   16:33   0:05 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_proxyserver -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-yarn -
Dyarn.log.dir=/var/log/hadoop-yarn -Dhadoop.log.file=yarn-yarn-
yarn      8646  1.0  1.2 1876916 385308 ?      Sl   16:33   0:46 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_resourcemanager -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-yarn
 -Dyarn.log.dir=/var/log/hadoop-yarn -Dhadoop.log.file=yarn-y
mapred    9265  0.2  0.8 1666628 260484 ?      Sl   16:33   0:12 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_historyserver -Xmx1000m -Dhadoop.log.dir=/usr/lib/hadoop/logs -
Dhadoop.log.file=hadoop.log -Dhadoop.home.dir=/usr/lib/hadoop
```

Vérification des dépendances à l'aide du référentiel d'artefacts d'Amazon EMR

Vous pouvez utiliser le référentiel d'artefacts Amazon EMR pour générer le code de travail Apache Hive et Apache Hadoop sur les versions exactes des bibliothèques et dépendances qui sont disponibles avec des versions Amazon EMR spécifiques, en commençant par la version 5.18.0 d'Amazon EMR . La création contre les artefacts Amazon EMR dans le référentiel contribue à éviter des problèmes de chemin d'accès de classe d'exécution en assurant que les versions des bibliothèques contre lesquelles la tâche est construite sont exactement les mêmes versions fournies au moment de l'exécution sur le cluster. Actuellement, les artefacts Amazon EMR ne sont disponibles que pour les builds Maven.

Pour accéder à l'artefact de référentiel, ajoutez l'URL du référentiel à votre fichier de paramètres Maven ou à un fichier de configuration `pom.xml` d'un projet spécifique. Vous pouvez alors spécifier les dépendances de votre configuration de projet. Pour les versions de dépendances, utilisez la version répertoriée sous Versions de composant pour la version souhaitée sur Versions Amazon EMR 5.x. Par exemple, les versions des composants pour la dernière version Amazon EMR sont disponibles à l'adresse the section called “Versions des composants 5.36.2”. Si un artefact pour votre projet n'est pas répertorié sous Versions de composant, spécifiez la version qui est répertoriée pour Hive et Hadoop dans cette version. Par exemple, pour les composants Hadoop dans la version 5.18.0 d'Amazon EMR, la version est `2.8.4-amzn-1`.

L'URL de référentiel d'artefact dispose de la syntaxe suivante :

```
https://s3-endpoint/region-ID-emr-artifacts/emr-release-label/repos/maven/
```

- *`s3-endpoint`*est le point de terminaison Amazon Simple Storage Service (Amazon S3) de la région du référentiel *`region-ID`* et la région correspondante. Par exemple : `s3.us-west-1.amazonaws.com` et `us-west-1`. Pour plus d'informations, consultez les points de terminaison Amazon S3 dans la Référence générale d'Amazon Web Services. Il n'y a pas de différence d'artefacts entre les régions. Cela vous permet de spécifier la région la plus pratique pour votre environnement de développement.
- *`emr-release-label`*est le label de publication du cluster Amazon EMR qui exécutera votre code. Les étiquettes de version respectent le format `emr-x.x.x`, par exemple, `emr-5.36.2`. Une série de versions EMR peut inclure plusieurs versions. Par exemple, si vous utilisez EMR version

5.24.1, utilisez la première étiquette de version EMR de la série 5.24, `emr-5.24.0`, dans l'URL du référentiel d'artefacts :

```
https://s3-endpoint/region-ID-emr-artifacts/emr-5.24.0/repos/maven/
```

Example Configuration pour Maven pom.xml

L'exemple de pom.xml ci-dessous configure un projet Maven pour se créer contre les artefacts Apache Hadoop et Apache Hive emr-5.18.0, à l'aide du référentiel d'artefacts dans us-west-1. Les versions d'instantanés ne sont pas disponibles dans le référentiel d'artefacts. Par conséquent, les instantanés sont désactivés dans le pom.xml. Les ellipses (`...`) dans l'exemple ci-dessous indiquent l'omission d'autres paramètres de configuration. Ne les copiez pas dans votre projet Maven.

```
<project>
 ...
 <repositories>
  ...
  <repository>
   <id>emr-5.18.0-artifacts</id>
   <name>EMR 5.18.0 Releases Repository</name>
   <releases>
    <enabled>true</enabled>
   </releases>
   <snapshots>
    <enabled>false</enabled>
   </snapshots>
   <url>https://s3.us-west-1.amazonaws.com/us-west-1-emr-artifacts/emr-5.18.0/repos/
maven/</url>
  </repository>
  ...
 </repositories>
 ...
 <dependencies>
  ...
  <dependency>
   <groupId>org.apache.hive</groupId>
   <artifactId>hive-exec</artifactId>
   <version>2.3.3-amzn-2</version>
  </dependency>
  <dependency>
   <groupId>org.apache.hadoop</groupId>
```

```
    <artifactId>hadoop-common</artifactId>
    <version>2.8.4-amzn-1</version>
   </dependency>
   ...
  </dependencies>

</project>
```

Système de fichiers EMR (EMRFS)

Le système de fichiers EMR (EMRFS) est une mise en œuvre de HDFS que tous les clusters Amazon EMR utilisent pour la lecture et l'écriture des fichiers réguliers depuis Amazon EMR directement dans Amazon S3. EMRFS permet de stocker des données persistantes dans Amazon S3 en vue de les utiliser avec Hadoop, tout en fournissant des fonctionnalités telles que le chiffrement des données.

Le chiffrement des données vous permet de chiffrer les objets qu'EMRFS écrit dans Amazon S3 et permet à EMRFS de travailler avec les données chiffrées dans Amazon S3. Si vous utilisez la version 4.8.0 ou ultérieure d'Amazon EMR, vous pouvez utiliser les configurations de sécurité pour configurer le chiffrement des objets EMRFS dans Amazon S3, ainsi que d'autres paramètres de chiffrement. Pour plus d'informations, consultez Options de chiffrement. Si vous utilisez une version précédente version de Amazon EMR, vous pouvez configurer manuellement les paramètres de chiffrement. Pour de plus amples informations, veuillez consulter Spécification du chiffrement Amazon S3 en utilisant les propriétés EMRFS.

Amazon S3 offre une forte cohérence entre lecture et écriture pour toutes les opérations GET, PUT et LIST dans l'ensemble Régions AWS. Cela signifie que ce que vous écrivez à l'aide d'EMRFS correspond à ce que vous allez lire sur Amazon S3, sans aucun impact sur les performances. Pour plus d'informations, consultez Modèle de cohérence de données Amazon S3.

Lorsque vous utilisez les versions 5.10.0 ou ultérieures d'Amazon EMR, vous pouvez utiliser différents rôles IAM pour les demandes EMRFS à Amazon S3 en fonction des utilisateurs du cluster, des groupes ou de l'emplacement des données EMRFS dans Amazon S3. Pour plus d'informations, consultez Configuration de rôles IAM pour les demandes EMRFS à Amazon S3.

> **Warning**
>
> Avant d'activer l'exécution spéculative pour les clusters Amazon EMR exécutant des tâches Apache Spark, veuillez consulter les informations suivantes.
> EMRFS inclut le validateur optimisé pour S3, une OutputCommitter implémentation optimisée pour écrire des fichiers sur Amazon S3 lors de l'utilisation d'EMRFS. Si vous activez la fonctionnalité d'exécution spéculative d'Apache Spark avec des applications qui écrivent des données sur Amazon S3 et n'utilisent pas le committer optimisé pour EMRFS S3, vous pouvez rencontrer les problèmes d'exactitude des données décrits dans le SPARK-10063. Cela peut se produire si vous utilisez des versions d'Amazon EMR antérieures à la version 5.19 d'Amazon EMR, ou si vous écrivez des fichiers dans Amazon S3 avec des formats tels

que ORC et CSV. Ces formats ne sont pas pris en charge par le validateur optimisé pour EMRFS S3. Pour une liste complète des exigences relatives à l'utilisation du transmetteur optimisé EMRFS S3, consultez Exigences relatives au transmetteur optimisé EMRFS S3. L'écriture directe EMRFS est généralement utilisée lorsque le validateur optimisé pour EMRFS S3 n'est pas pris en charge, par exemple lors de l'écriture de ce qui suit :

- Un format de sortie autre que Parquet, tel que ORC ou texte.
- Fichiers Hadoop à l'aide de l'API Spark RDD.
- Parquet utilisant Hive. SerDe Consultez Conversion de la table Hive Metastore Parquet.

L'écriture directe EMRFS n'est pas utilisée dans les scénarios suivants :

- Lorsque le validateur EMRFS optimisé pour S3 est activé. Consultez Exigences du valideur EMRFS optimisé pour S3.
- Lors de l'écriture de partitions dynamiques partitionOverwriteMode avec la valeur dynamic.
- Lorsque vous écrivez sur des emplacements de partition personnalisés, tels que des emplacements non conformes à la convention d'emplacement de partition par défaut de Hive.
- Lorsque vous utilisez des systèmes de fichiers autres qu'EMRFS, tels que l'écriture sur HDFS ou l'utilisation du système de fichiers S3A.

Pour déterminer si votre application utilise l'écriture directe dans Amazon EMR 5.14.0 ou version ultérieure, activez la journalisation Spark INFO. Si une ligne de journal contenant le texte « Direct Write : ENABLED » est présente dans les journaux du pilote Spark ou dans les journaux des conteneurs de l'exécuteur Spark, cela signifie que votre application Spark a écrit en écriture directe.
Par défaut, l'exécution spéculative est activée `OFF` sur Amazon EMRclusters. Nous vous recommandons vivement de ne pas activer l'exécution spéculative si l'une ou l'autre de ces conditions est remplie :

- Vous écrivez des données sur Amazon S3.
- Les données sont écrites dans un format autre qu'Apache Parquet ou dans le format Apache Parquet sans utiliser le validateur optimisé pour EMRFS S3.

Si vous activez l'exécution spéculative de Spark et que vous écrivez des données sur Amazon S3 à l'aide de l'écriture directe EMRFS, vous risquez de subir des pertes de données intermittentes. Lorsque vous écrivez des données dans HDFS ou que vous écrivez des données dans Parquet à l'aide du validateur optimisé EMRFS S3, Amazon EMR n'utilise pas l'écriture directe et ce problème ne se produit pas.
Si vous devez écrire des données dans des formats utilisant l'écriture directe EMRFS depuis Spark vers Amazon S3 et utiliser l'exécution spéculative, nous vous recommandons d'écrire sur HDFS, puis de transférer les fichiers de sortie vers Amazon S3 à l'aide de S3DistCp.

Rubriques

- Vue cohérente
- Autorisation d'accès aux données EMRFS dans Amazon S3
- Gestion du point de AWS Security Token Service terminaison par défaut
- Spécification du chiffrement Amazon S3 en utilisant les propriétés EMRFS

Vue cohérente

⚠ Warning

Le 1er juin 2023, la vue cohérente d'EMRFS atteindra la fin du support standard pour les futures versions d'Amazon EMR. La vue cohérente EMRFS continuera de fonctionner pour les versions existantes.

Avec le lancement de la forte read-after-write cohérence d'Amazon S3 le 1er décembre 2020, vous n'avez plus besoin d'utiliser la vue cohérente EMRFS (EMRFS CV) avec vos clusters Amazon EMR. EMRFS CV est une fonctionnalité optionnelle qui permet aux clusters Amazon EMR de vérifier la cohérence des listes et la cohérence des objets read-after-write Amazon S3. Lorsque vous créez un cluster et que le CV EMRFS est activé, Amazon EMR crée une base de données Amazon DynamoDB pour stocker les métadonnées des objets qu'il utilise pour suivre la liste et la cohérence des objets S3. read-after-write Vous pouvez désormais désactiver EMRFS CV et supprimer la base de données DynamoDB qu'il utilise afin de ne pas générer de coûts supplémentaires. Les procédures

suivantes expliquent comment vérifier la fonctionnalité CV, la désactiver et supprimer la base de données DynamoDB qu'elle utilise.

Pour vérifier si vous utilisez la fonction EMRFS CV

1. Accédez à l'onglet Configuration. Si votre cluster possède la configuration suivante, il utilise EMRFS CV.

   ```
   Classification=emrfs-site,Property=fs.s3.consistent,Value=true
   ```

2. Vous pouvez également utiliser le AWS CLI pour décrire votre cluster avec l'`describe-cluster`API. Si le résultat en contient `fs.s3.consistent: true`, votre cluster utilise EMRFS CV.

Pour désactiver le CV EMRFS sur vos clusters Amazon EMR

Pour désactiver la fonction EMRFS CV, utilisez l'une des trois options suivantes. Vous devez tester ces options dans votre environnement de test avant de les appliquer à vos environnements de production.

1. Pour arrêter votre cluster existant et démarrer un nouveau cluster sans les options EMRFS CV.

   a. Avant d'arrêter votre cluster, assurez-vous de sauvegarder vos données et d'en informer les utilisateurs.

   b. Pour arrêter votre cluster, suivez les instructions de la section Résilier un cluster.

   c. Si vous utilisez la console Amazon EMR pour créer un nouveau cluster, accédez aux Options avancées. Dans la section Modifier les paramètres du logiciel, désélectionnez l'option permettant d'activer EMRFS CV. Si la case à cocher pour une vue cohérente EMRFS est disponible, désactivez-la.

   d. Si vous utilisez AWS CLI pour créer un nouveau cluster avec l'`create-cluster`API, n'utilisez pas l'`--emrfs`option qui active EMRFS CV.

   e. Si vous utilisez un SDK ou AWS CloudFormation pour créer un nouveau cluster, n'utilisez aucune des configurations répertoriées dans Configurer une vue cohérente.

2. Pour cloner un cluster et supprimer le CV EMRFS

   a. Dans la console Amazon EMR, choisissez le cluster qui utilise EMRFS CV.

   b. En haut de la page Détails du cluster, sélectionnez Cloner.

c. Choisissez Précédent et passez à l'Étape 1 : Logiciel et étapes.

d. Dans Modifier les paramètres du logiciel, supprimez le EMRFS CV. Dans Modifier la configuration, supprimez les configurations suivantes dans la classification `emrfs-site`. Si vous chargez du JSON à partir d'un compartiment S3, vous devez modifier votre objet S3.

```
[
 {"classification":
  "emrfs-site",
  "properties": {
   "fs.s3.consistent.retryPeriodSeconds":"10",
   "fs.s3.consistent":"true",
   "fs.s3.consistent.retryCount":"5",
   "fs.s3.consistent.metadata.tableName":"EmrFSMetadata"
  }
 }
]
```

3. Pour supprimer le CV EMRFS d'un cluster qui utilise des groupes d'instances

a. Utilisez la commande suivante pour vérifier si un seul cluster EMR utilise la table DynamoDB associée à EMRFS CV, ou si plusieurs clusters partagent la table. Le nom de la table est spécifié dans `fs.s3.consistent.metadata.tableName`, comme décrit dans Configuration de la vue cohérente. Le nom de table par défaut utilisé par EMRFS CV est `EmrFSMetadata`.

```
aws emr describe-cluster --cluster-id j-XXXXX | grep
 fs.s3.consistent.metadata.tableName
```

b. Si votre cluster ne partage pas votre base de données DynamoDB avec un autre cluster, utilisez la commande suivante pour reconfigurer le cluster et désactiver EMRFS CV. Pour plus d'informations, consultez Reconfiguration d'un groupe d'instances dans un cluster en cours d'exécution.

```
aws emr modify-instance-groups --cli-input-json file://disable-emrfs-1.json
```

Cette commande ouvre le fichier que vous souhaitez modifier. Modifiez le fichier avec l'une des configurations suivantes.

```
{
 "ClusterId": "j-xxxx",
```

```
 "InstanceGroups": [
  {
   "InstanceGroupId": "ig-xxxx",
   "Configurations": [
    {
     "Classification": "emrfs-site",
     "Properties": {
      "fs.s3.consistent": "false"
     },
     "Configurations": []
    }
   ]
  }
 ]
}
```

c. Si votre cluster partage la table DynamoDB avec un autre cluster, désactivez EMRFS CV sur tous les clusters à un moment où aucun cluster ne modifie les objets de l'emplacement S3 partagé.

Pour supprimer les ressources Amazon DynamoDB associées au CV EMRFS

Après avoir supprimé le CV EMRFS de vos clusters Amazon EMR, supprimez les ressources DynamoDB associées au CV EMRFS. Jusqu'à ce que vous le fassiez, vous continuez à payer des frais DynamoDB associés à EMRFS CV.

1. Vérifiez les CloudWatch métriques de votre table DynamoDB et vérifiez que la table n'est utilisée par aucun cluster.
2. Supprimez la table DynamoDB.

```
aws dynamodb delete-table --table-name <your-table-name>
```

Pour supprimer les ressources Amazon SQS associées au CV EMRFS

1. Si vous avez configuré votre cluster pour envoyer des notifications d'incohérence à Amazon SQS, vous pouvez supprimer toutes les files d'attente SQS.
2. Recherchez le nom de file d'attente Amazon SQS spécifié dans `fs.s3.consistent.notification.SQS.queueName`, comme décrit dans Configurer une

vue cohérente. Le format de nom de file d'attente par défaut est `EMRFS-Inconsistency-<j-cluster ID>`.

```
aws sqs list-queues | grep 'EMRFS-Inconsistency'
aws sqs delete-queue -queue-url <your-queue-url>
```

Pour arrêter d'utiliser l'interface de ligne de commande EMRFS

- La CLI EMRFS gère les métadonnées générées par EMRFS CV. La prise en charge standard d'EMRFS CV arrivant à son terme dans les futures versions d'Amazon EMR, la prise en charge de la CLI EMRFS atteindra également sa fin.

Rubriques

- Activation de la vue cohérente
- Comprendre comment la vue cohérente d'EMRFS effectue le suivi des objets dans Amazon S3
- Logique des nouvelles tentatives
- Métadonnées de la vue cohérente EMRFS
- Configuration des notifications de cohérence pour CloudWatch et Amazon SQS
- Configuration de la vue cohérente
- Référence des commandes CLI EMRFS

Activation de la vue cohérente

Vous pouvez activer le chiffrement côté serveur Amazon S3 ou la vue cohérente pour EMRFS en utilisant la AWS Management Console AWS CLI, ou la classification de configuration. `emrfs-site`

Pour configurer la vue cohérente à l'aide de la console

1. Accédez à la nouvelle console Amazon EMR et sélectionnez Changer pour l'ancienne console depuis le menu latéral. Pour plus d'informations sur ce qu'implique le passage à l'ancienne console, consultez la rubrique Utilisation de l'ancienne console.
2. Choisissez Créer un cluster et Go to advanced options (Aller aux options avancées).
3. Choisissez des valeurs pour Step 1: Software and Steps (Étape 1 : Logiciel et étapes) et Step 2: Hardware (Étape 2 : Matériel).

4. Pour Step 3: General Cluster Settings (Étape 3 : Paramètres généraux de cluster), sous Additional Options (Options supplémentaires), choisissez EMRFS consistent view (Vue cohérente EMRFS).
5. Pour EMRFS Metadata store (Boutique de métadonnées EMRFS), saisissez le nom de votre magasin de métadonnées. La valeur par défaut est **`EmrFSMetadata`**. Si la FSMetadata table Emr n'existe pas, elle est créée pour vous dans DynamoDB.

   > **Note**
   >
   > Amazon EMR ne supprime pas automatiquement les métadonnées EMRFS de DynamoDB lorsque le cluster est résilié.

6. Pour Number of retries (Nombre de nouvelles tentatives), tapez une valeur entière. Si une incohérence est détectée, EMRFS essaie d'appeler Amazon S3 ce nombre de fois. La valeur par défaut est **`5`**.
7. Pour Retry period (in seconds) (Période de nouvelle tentative (en secondes)), tapez une valeur entière. Il s'agit du temps pendant lequel EMRFS attend entre les nouvelles tentatives. La valeur par défaut est **`10`**.

   > **Note**
   >
   > Les nouvelles tentatives ultérieures utilisent une interruption exponentielle.

Lancer un cluster avec une vue cohérente activée à l'aide de l' AWS CLI

Nous vous recommandons d'installer la version actuelle de AWS CLI. Pour télécharger la dernière version, voir https://aws.amazon.com/cli/.

- > **Note**
  >
  > Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

  ```
  aws emr create-cluster --instance-type m5.xlarge --instance-count 3 --emrfs
   Consistent=true \
  ```

```
--release-label emr-7.9.0 --ec2-attributes KeyName=myKey
```

Pour vérifier si la vue cohérente est activée à l'aide du AWS Management Console

- Pour vérifier si la vue cohérente est activée dans la console, accédez à la Liste de clusters et sélectionnez le nom de votre cluster pour afficher Cluster Details (Détails de cluster). Le champ « EMRFS consistent view » a la valeur `Enabled` ou `Disabled`.

Pour vérifier si la vue cohérente est activée en examinant le fichier **`emrfs-site.xml`**

- Vous pouvez vérifier si la cohérence est activée en inspectant le fichier de configuration `emrfs-site.xml` sur le nœud maître du cluster. Si la valeur booléenne pour `fs.s3.consistent` est définie sur `true`, la vue cohérente est activée pour les opérations du système de fichiers impliquant Amazon S3.

Comprendre comment la vue cohérente d'EMRFS effectue le suivi des objets dans Amazon S3

EMRFS crée une vue cohérente des objets dans Amazon S3 en ajoutant des informations sur ces objets aux métadonnées EMRFS. EMRFS ajoute ces listes à ses métadonnées lorsque :

- Objet écrit par EMRFS au cours d'une tâche Amazon EMR.
- un objet est synchronisé avec les métadonnées EMRFS ou importé dans les métadonnées EMRFS à l'aide de l'interface de ligne de commande d'EMRFS.

Les objets lus par EMRFS ne sont pas automatiquement ajoutés aux métadonnées. Lorsqu'EMRFS supprime un objet, une liste demeure dans les métadonnées avec un état supprimé jusqu'à ce que cette liste soit purgée à l'aide de l'interface de ligne de commande d'EMRFS. Pour plus d'informations sur l'interface de ligne de commande, consultez Référence des commandes CLI EMRFS. Pour plus d'informations sur la purge des listes dans les métadonnées EMRFS, consultez Métadonnées de la vue cohérente EMRFS.

Pour chaque opération Amazon S3, EMRFS vérifie les métadonnées pour obtenir des informations sur l'ensemble des objets dans la vue cohérente. Si EMRFS considère que Amazon S3 est incohérent au cours d'une de ces opérations, il réessaie l'opération selon les paramètres définis dans les propriétés de configuration `emrfs-site`. Une fois qu'EMRFS a épuisé les tentatives, il lève une

exception `ConsistencyException` ou consigne l'exception et continue le flux de travail. Pour plus d'informations sur la logique des nouvelles tentatives, consultez Logique des nouvelles tentatives. Vous pouvez trouver des exceptions `ConsistencyExceptions` dans vos journaux, par exemple :

- listStatus : aucun objet Amazon S3 pour l'élément de métadonnées `/S3_bucket/dir/object`
- getFileStatus: la clé `dir/file` est présente dans les métadonnées, mais pas dans Amazon S3

Si vous supprimez un objet qui est suivi dans la vue cohérente d'EMRFS directement à partir d'Amazon S3 , EMRFS traite cet objet comme incohérent, car il reste répertorié dans les métadonnées tel qu'il figure dans Amazon S3. Si vos métadonnées cessent d'être synchronisées avec les objets qu'EMRFS suit dans Amazon S3, vous pouvez utiliser la sous-commande sync sur l'interface de ligne de commande d'EMRFS pour réinitialiser les métadonnées afin qu'elles tiennent compte d'Amazon S3. Pour découvrir les différences entre les métadonnées et Amazon S3, utilisez l'outil diff. Enfin, EMRFS a uniquement une vue cohérente des objets référencés dans les métadonnées. D'autres objets peuvent figurer dans le même chemin d'accès Amazon S3 et ne pas être suivis. Quand EMRFS répertorie les objets dans un chemin d'accès Amazon S3, il renvoie le sur-ensemble des objets suivis dans les métadonnées et de ceux dans ce chemin d'accès Amazon S3.

Logique des nouvelles tentatives

EMRFS essaie de vérifier la cohérence des listes pour les objets suivis dans ses métadonnées pour un certain nombre de tentatives. La valeur par défaut est 5. Au cas où le nombre de nouvelles tentatives est dépassé, la tâche initiale retourne un échec, à moins que `fs.s3.consistent.throwExceptionOnInconsistency` ait la valeur `false`, auquel cas elle consignera uniquement les objets suivis comme incohérents. EMRFS utilise une stratégie de nouvelles tentatives d'interruption exponentielle par défaut, mais vous pouvez également la configurer comme une stratégie fixe. Les utilisateurs peuvent également réessayer pendant un certain temps avant de passer au reste de leur tâche sans lever d'exception. Ils peuvent y parvenir en définissant `fs.s3.consistent.throwExceptionOnInconsistency` sur `false`, `fs.s3.consistent.retryPolicyType` sur `fixed` et `fs.s3.consistent.retryPeriodSeconds` sur la valeur de leur choix. L'exemple suivant crée un cluster avec la cohérence activée, qui consigne les incohérences et définit un intervalle fixe de nouvelle tentative de 10 secondes :

Example Configuration de la période de nouvelle tentative sur une valeur fixe

```
aws emr create-cluster --release-label emr-7.9.0 \
--instance-type m5.xlarge --instance-count 1 \
```

```
--emrfs Consistent=true,Args=[fs.s3.consistent.throwExceptionOnInconsistency=false,
 fs.s3.consistent.retryPolicyType=fixed,fs.s3.consistent.retryPeriodSeconds=10] --ec2-
attributes KeyName=myKey
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

Pour de plus amples informations, veuillez consulter Vue cohérente.

Les configurations EMRFS pour IMDS reçoivent des appels régionaux

EMRFS s'appuie sur l'IMDS (service de métadonnées d'instance) pour obtenir la région de l'instance et Amazon S3, DynamoDB ou les points de terminaison. AWS KMS Cependant, l'IMDS a une limite quant au nombre de demandes qu'il peut traiter, et les demandes qui dépassent cette limite échoueront. Cette limite IMDS peut entraîner l'échec de l'initialisation d'EMRFS et l'échec de la requête ou de la commande. Vous pouvez utiliser le mécanisme de backoff exponentiel aléatoire suivant et les propriétés de configuration d'une région de secours dans le fichier emrfs-site.xml pour résoudre le scénario dans lequel toutes les tentatives échouent.

```
<property>
    <name>fs.s3.region.retryCount</name>
    <value>3</value>
    <description>
    Maximum retries that would be attempted to get AWS region.
    </description>
</property>
<property>
    <name>fs.s3.region.retryPeriodSeconds</name>
    <value>3</value>
    <description>
    Base sleep time in second for each get-region retry.
    </description>
</property>
<property>
    <name>fs.s3.region.fallback</name>
    <value>us-east-1</value>
    <description>
```

```
    Fallback to this region after maximum retries for getting AWS region have been
 reached.
    </description>
</property>
```

Métadonnées de la vue cohérente EMRFS

La vue cohérente d'EMRFS assure la cohérence à l'aide d'une table DynamoDB pour suivre les objets dans Amazon S3 qui ont été synchronisés avec ou créés par le système EMRFS. Les métadonnées sont utilisées pour suivre toutes les opérations (lecture, écriture, mise à jour et copie) et aucun contenu réel n'y est stocké. Ces métadonnées sont utilisées pour vérifier si les objets ou les métadonnées provenant d'Amazon S3 correspondent à ce qui est prévu. Cette confirmation permet à EMRFS de contrôler la cohérence des listes et read-after-write la cohérence des nouveaux objets qu'EMRFS écrit sur Amazon S3 ou pour les objets synchronisés avec EMRFS. Plusieurs clusters peuvent partager les mêmes métadonnées.

Comment ajouter des entrées aux métadonnées

Vous pouvez utiliser les sous-commandes `sync` ou `import` pour ajouter des entrées aux métadonnées. `sync` reflète l'état des objets Amazon S3 dans un chemin d'accès lorsque la sous-commande `import` est utilisée strictement pour ajouter de nouvelles entrées aux métadonnées. Pour de plus amples informations, veuillez consulter Référence des commandes CLI EMRFS.

Vérification des différences entre les métadonnées et les objets dans Amazon S3

Pour vérifier les différences entre les métadonnées et Amazon S3, utilisez la sous-commande `diff` de l'interface de ligne de commande d'EMRFS. Pour de plus amples informations, veuillez consulter Référence des commandes CLI EMRFS.

Comment savoir si les opérations de métadonnées sont limitées

EMRFS définit les limites de capacité de débit par défaut sur les métadonnées pour ses opérations de lecture et d'écriture à 500 et 100 unités, respectivement. Un grand nombre d'objets ou de compartiments peut amener les opérations à dépasser cette capacité, auquel cas DynamoDB limitera les opérations. Par exemple, une application peut amener EMRFS à lever une exception `ProvisionedThroughputExceededException` si vous effectuez une opération qui dépasse ces limites de capacité. Lors d'une limitation, l'outil d'interface de ligne de commande d'EMRFS tente de réécrire dans la table DynamoDB à l'aide d'un backoff exponentiel jusqu'à ce que l'opération s'achève ou qu'il atteigne la valeur maximale de nouvelles tentatives d'écriture d'objets d'Amazon EMR sur Amazon S3.

Vous pouvez configurer vos propres limites de capacité de débit. Toutefois, DynamoDB impose des limites de partition strictes de 3 000 unités de capacité de lecture RCUs () et 1 000 unités de capacité d'écriture WCUs () par seconde pour les opérations de lecture et d'écriture. Pour éviter les `sync` défaillances causées par la régulation, nous vous recommandons de limiter le débit des opérations de lecture à moins de 3 000 RCUs et des opérations d'écriture à moins de 1 000. WCUs Pour obtenir des instructions sur la définition de limites de capacité de débit personnalisées, consultez Configuration de la vue cohérente.

Vous pouvez également consulter CloudWatch les métriques Amazon pour vos métadonnées EMRFS dans la console DynamoDB, qui vous indiquent le nombre de demandes de lecture et d'écriture limitées. Si le nombre de demandes limitées n'est pas nul, votre application peut éventuellement bénéficier d'une capacité grandissante de débit allouée pour les opérations de lecture ou d'écriture. Vous pouvez également réaliser une amélioration des performances si vous voyez que vos opérations approchent de la capacité de débit allouée maximale pour les lectures ou les écritures pendant une période de temps prolongée.

Caractéristiques de débit pour les opérations EMRFS notables

La valeur par défaut pour les opérations de lecture et d'écriture est de 400 et de 100 unités de capacité de débit, respectivement. Les caractéristiques de performance suivantes vous donnent une idée du débit requis pour certaines opérations. Ces tests ont été effectués à l'aide d'un cluster `m3.large` à un seul nœud. Toutes les opérations étaient à thread unique. Les performances varient considérablement en fonction des caractéristiques d'application spécifiques et il peut être nécessaire d'expérimenter afin d'optimiser les opérations de système de fichiers.

Opération	Moyenne read-per-second	Moyenne write-per-second
create (objet)	26,79	6,70
delete (objet)	10,79	10,79
delete (répertoire contenant 1 000 objets)	21,79	338,40
getFileStatus(objet)	34,70	0
getFileStatus(annuaire)	19,96	0

Opération	Moyenne read-per-second	Moyenne write-per-second
listStatus (répertoire contenant 1 objet)	43,31	0
listStatus (répertoire contenant 10 objets)	44,34	0
listStatus (répertoire contenant 100 objets)	84,44	0
listStatus (répertoire contenant 1 000 objets)	308,81	0
listStatus (répertoire contenant 10 000 objets)	416,05	0
listStatus (répertoire contenant 100 000 objets)	823,56	0
listStatus (répertoire contenant 1 million d'objets)	882,36	0
mkdir (en continu pendant 120 secondes)	24,18	4,03
mkdir	12,59	0
rename (objet)	19,53	4,88
rename (répertoire contenant 1 000 objets)	23,22	339,34

Pour soumettre une étape qui purge les anciennes données à partir de votre magasin de métadonnées

Les utilisateurs peuvent supprimer des entrées particulières dans les métadonnées basées sur DynamoDB. Cela peut aider à réduire les coûts de stockage associés à la table. Les utilisateurs

ont la possibilité de purger manuellement ou par programmation des entrées particulières à l'aide de la sous-commande `delete` de l'interface de ligne de commande d'EMRFS. Toutefois, si vous supprimez des entrées à partir des métadonnées, EMRFS n'effectue plus aucun contrôle de cohérence.

La purge par programmation à l'issue d'une tâche peut être effectuée en soumettant une étape finale à votre cluster qui exécute une commande sur l'interface de ligne de commande d'EMRFS. Par exemple, tapez la commande suivante pour soumettre une étape à votre cluster afin de supprimer toutes les entrées de plus de deux jours.

```
aws emr add-steps --cluster-id j-2AL4XXXXXX5T9 --steps Name="emrfsCLI",Jar="command-
runner.jar",Args=["emrfs","delete","--time","2","--time-unit","days"]
{
    "StepIds": [
        "s-B12345678902"
    ]
}
```

Utilisez la StepId valeur renvoyée pour vérifier les journaux concernant le résultat de l'opération.

Configuration des notifications de cohérence pour CloudWatch et Amazon SQS

Vous pouvez activer CloudWatch les métriques et les messages Amazon SQS dans EMRFS pour les problèmes de cohérence éventuels d'Amazon S3.

CloudWatch

Lorsque CloudWatch les métriques sont activées, une métrique nommée Incohérence est poussée chaque fois qu'un appel d'`FileSystem`API échoue en raison de la cohérence à terme d'Amazon S3.

Pour afficher CloudWatch les métriques concernant les problèmes de cohérence éventuels d'Amazon S3

Pour afficher la métrique d'incohérence dans la CloudWatch console, sélectionnez les métriques EMRFS, puis sélectionnez une paire JobFlowId/Metric Name. Par exemple : `j-162XXXXXXM2CU ListStatus`, `j-162XXXXXXM2CU GetFileStatus`, etc.

1. Ouvrez la CloudWatch console à l'adresse https://console.aws.amazon.com/cloudwatch/.

2. Dans le Tableau de bord, dans la section Metrics (Métriques), choisissez EMRFS.
3. Dans le volet Job Flow Metrics, sélectionnez une ou plusieurs paires JobFlowId/Metric Name. Une représentation graphique des métriques s'affiche dans la fenêtre ci-dessous.

Amazon SQS

Lorsque les notifications Amazon SQS sont activées, une file d'attente Amazon SQS portant le nom `EMRFS-Inconsistency-<jobFlowId>` est créée lors de l'initialisation d'EMRFS. Les messages Amazon SQS sont placés dans la file d'attente lorsqu'un appel d'API `FileSystem` échoue en raison de la cohérence éventuelle d'Amazon S3. Le message contient des informations telles que l'API JobFlowId, une liste de chemins incohérents, une trace de pile, etc. Les messages peuvent être lus à l'aide de la console Amazon SQS ou de la commande `read-sqs` d'EMRFS.

Pour gérer les messages Amazon SQS en cas de problèmes de cohérence éventuels d'Amazon S3

Les messages Amazon SQS pour les problèmes de cohérence éventuelle d'Amazon S3 peuvent être lus à l'aide de l'interface de ligne de commande d'EMRFS. Pour lire les messages à partir d'une file d'attente Amazon SQS d'EMRFS, tapez la commande `read-sqs` et spécifiez un emplacement de sortie dans le système de fichiers local du nœud principal pour le fichier de sortie obtenu.

Vous pouvez également supprimer une file d'attente Amazon SQS d'EMRFS à l'aide de la commande `delete-sqs`.

1. Pour lire les messages à partir d'une file d'attente Amazon SQS, tapez la commande suivante. Remplacez *queuename* par le nom de la file d'attente Amazon SQS que vous avez configuré et remplacez par le nom de la file d'attente Amazon SQS que vous avez configuré et remplacez */path/filename* par le chemin d'accès du fichier de sortie :

   ```
   emrfs read-sqs --queue-name queuename --output-file /path/filename
   ```

   Par exemple, pour lire et générer des messages Amazon SQS à partir de la file d'attente par défaut, tapez :

   ```
   emrfs read-sqs --queue-name EMRFS-Inconsistency-j-162XXXXXXM2CU --output-file /
   path/filename
   ```

> **Note**
>
> Vous pouvez également utiliser les raccourcis `-q` et `-o` à la place de `--queue-name` et `--output-file` respectivement.

2. Pour supprimer une file d'attente Amazon SQS, tapez la commande suivante :

```
emrfs delete-sqs --queue-name queuename
```

Par exemple, pour supprimer la file d'attente par défaut, tapez :

```
emrfs delete-sqs --queue-name EMRFS-Inconsistency-j-162XXXXXXM2CU
```

> **Note**
>
> Vous pouvez également utiliser le raccourci `-q` à la place de `--queue-name`.

Configuration de la vue cohérente

Vous pouvez configurer d'autres paramètres pour la vue cohérente en les fournissant à l'aide des propriétés de configuration pour les propriétés `emrfs-site`. Par exemple, vous pouvez choisir un autre débit DynamoDB par défaut en fournissant les arguments suivants à l'interface de ligne de commande, à l'aide de l'option `--emrfs`, en utilisant la classification de configuration de site emrfs (versions 4.x et ultérieures d'Amazon EMR uniquement), ou une action d'amorçage pour configurer le fichier emrfs-site.xml sur le nœud principal :

Example Modification des valeurs de lecture et d'écriture de métadonnées par défaut au lancement du cluster

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge \
--emrfs Consistent=true,Args=[fs.s3.consistent.metadata.read.capacity=600,\
fs.s3.consistent.metadata.write.capacity=300] --ec2-attributes KeyName=myKey
```

Vous pouvez également utiliser le fichier de configuration suivant et l'enregistrer localement ou dans Amazon S3 :

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "fs.s3.consistent.metadata.read.capacity": "600",
        "fs.s3.consistent.metadata.write.capacity": "300"
      }
    }
 ]
```

Utilisez la configuration que vous avez créée avec la syntaxe suivante :

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Hive \
--instance-type m5.xlarge --instance-count 2 --configurations file://./myConfig.json
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

Les options suivantes peuvent être définies à l'aide de configurations ou d' AWS CLI `--emrfs`arguments. Pour plus d'informations sur ces arguments, consultez la Référence des commandes AWS CLI.

Propriétés **`emrfs-site.xml`** pour la vue cohérente

Propriété	Valeur par défaut	Description
`fs.s3.consistent`	**`false`**	Lorsque la valeur est **`true`**, cette propriété configure EMRFS pour utiliser DynamoDB afin d'assurer la cohérence.
`fs.s3.consistent.retryPolicyType`	**`exponential`**	Cette propriété identifie la stratégie à utiliser lors d'une nouvelle tentative pour des

Propriété	Valeur par défaut	Description
		problèmes de cohérence. Les options comprennent : exponential, fixed et none.
`fs.s3.consistent.retryPeriodSeconds`	**1**	Cette propriété définit la durée d'attente entre les tentatives de relance de cohérence.
`fs.s3.consistent.retryCount`	**10**	Cette propriété définit le nombre maximal de nouvelles tentatives lorsqu'une incohérence est détectée.
`fs.s3.consistent.throwExceptionOnInconsistency`	**true**	Cette propriété détermine s'il convient de lever ou de consigner une exception de cohérence. Lorsque la valeur est **true**, une exception `ConsistencyException` est levée.
`fs.s3.consistent.metadata.autoCreate`	**true**	Lorsque la valeur est **true**, cette propriété permet la création automatique de tables de métadonnées.

Propriété	Valeur par défaut	Description
`fs.s3.consistent.metadata.etag.verification.enabled`	**`true`**	Avec Amazon EMR 5.29.0, cette propriété est activée par défaut. Lorsque l'option est activée, EMRFS utilise S3 ETags pour vérifier que les objets en cours de lecture correspondent à la dernière version disponible. Cette fonctionnalité est utile pour les cas read-after-update d'utilisation dans lesquels des fichiers sur S3 sont remplacés tout en conservant le même nom. Cette fonctionnalité de ETag vérification ne fonctionne pas actuellement avec S3 Select.
`fs.s3.consistent.metadata.tableName`	**`EmrFSMetadata`**	Cette propriété spécifie le nom de la table de métadonnées dans DynamoDB.
`fs.s3.consistent.metadata.read.capacity`	**`500`**	Cette propriété spécifie la capacité de lecture DynamoDB à mettre en service lorsque la table de métadonnées est créée.
`fs.s3.consistent.metadata.write.capacity`	**`100`**	Cette propriété spécifie la capacité d'écriture DynamoDB à mettre en service lorsque la table de métadonnées est créée.

Propriété	Valeur par défaut	Description
`fs.s3.consistent.fastList`	**`true`**	Lorsque la valeur est **`true`**, cette propriété utilise plusieurs fils pour répertorier un répertoire (si nécessaire). La cohérence doit être activée pour pouvoir utiliser cette propriété.
`fs.s3.consistent.fastList.prefetchMetadata`	**`false`**	Lorsque la valeur est **`true`**, cette propriété permet l'extraction préalable de métadonnées pour les répertoires contenant plus de 20 000 éléments.
`fs.s3.consistent.notification.CloudWatch`	**`false`**	Lorsque la valeur est définie sur**`true`**, les CloudWatch métriques sont activées pour les appels d' FileSystem API qui échouent en raison de problèmes de cohérence à terme d'Amazon S3.
`fs.s3.consistent.notification.SQS`	**`false`**	Lorsque la valeur est **`true`**, les notifications de cohérence éventuelle font l'objet d'un push vers une file d'attente Amazon SQS.
`fs.s3.consistent.notification.SQS.queueName`	**`EMRFS-Inconsistency-<jobFlowId>`**	La modification de cette propriété vous permet de spécifier votre propre nom de file d'attente SQS pour les messages concernant les problèmes de cohérence à terme d'Amazon S3.

Propriété	Valeur par défaut	Description
`fs.s3.consistent.notification.SQS.customMsg`	**`none`**	Cette propriété vous permet de spécifier des informations personnalisées incluses dans les messages SQS concernant les problèmes de cohérence à terme d'Amazon S3. Si une valeur n'est pas spécifiée pour cette propriété, le champ correspondant dans le message est vide.
`fs.s3.consistent.dynamodb.endpoint`	**`none`**	Cette propriété vous permet de spécifier un point de terminaison DynamoDB personnalisé pour vos métadonnées de vue cohérente.
`fs.s3.useRequesterPaysHeader`	**`false`**	Lorsqu'elle est définie sur **`true`**, cette propriété autorise les requêtes Amazon S3 à être envoyées dans des compartiments avec l'option de payeur de demandes activée.

Référence des commandes CLI EMRFS

L'interface de ligne de commande EMRFS est installée par défaut sur tous les nœuds principaux du cluster créé à l'aide de Amazon EMR 3.2.1 ou version supérieure. Vous pouvez utiliser l'interface de ligne de commande EMRFS pour gérer les métadonnées pour la vue cohérente.

> **Note**
>
> La emrfs commande n'est prise en charge qu'avec l'émulation du terminal VT1 00. Cependant, elle peut fonctionner avec les autres modes d'émulateur de terminal.

Commande emrfs de niveau supérieur

La commande emrfs de niveau supérieur prend en charge la structure suivante.

```
emrfs [describe-metadata | set-metadata-capacity | delete-metadata | create-metadata |
 \
list-metadata-stores | diff | delete | sync | import ] [options] [arguments]
```

Spécifiez des [options], avec ou sans arguments, comme décrit dans le tableau suivant. Pour les [options] spécifiques à des sous-commandes (`describe-metadata`, `set-metadata-capacity`, etc.), consultez chaque sous-commande ci-dessous.

[Options] pour emrfs

Option	Description	Obligatoire
`-a` *AWS_ACCESS_KEY_ID* | `--access-key` *AWS_ACCESS_KEY_ID*	La clé AWS d'accès que vous utilisez pour écrire des objets sur Amazon S3 et pour créer un magasin de métadonnées ou y accéder dans DynamoDB. Par défaut, *AWS_ACCESS_KEY_ID* est défini sur la clé d'accès utilisée pour créer le cluster.	Non
`-s` *AWS_SECRET_ACCESS_KEY* | `--secret-key` *AWS_SECRET_ACCESS_KEY*	La clé AWS secrète associée à la clé d'accès que vous utilisez pour écrire des objets sur Amazon S3 et pour créer un magasin de métadonnées ou y accéder dans DynamoDB. Par défaut, *AWS_SECRET_ACCESS_KEY* est défini sur la clé secrète associée à la clé d'accès utilisée pour créer le cluster.	Non
`-v` | `--verbose`	Détaillez le résultat.	Non
`-h` | `--help`	Affiche le message d'aide pour la commande `emrfs` avec une instruction d'utilisation.	Non

Sous commande describe-metadata emrfs

[Options] pour emrfs describe-metadata

Option	Description	Obligatoire
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* est le nom de la table de métadonnées DynamoDB. Si l'*METADATA_NAME* argument n'est pas fourni, la valeur par défaut est`EmrFSMetadata` .	Non

Example Exemple describe-metadata emrfs

L'exemple suivant décrit la table de métadonnées par défaut.

```
$ emrfs describe-metadata
EmrFSMetadata
  read-capacity: 400
  write-capacity: 100
  status: ACTIVE
  approximate-item-count (6 hour delay): 12
```

sous-commande emrfs set-metadata-capacity

[Options] pour emrfs set-metadata-capacity

Option	Description	Obligatoire
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* est le nom de la table de métadonnées DynamoDB. Si l'*METADATA_NAME* argument n'est pas fourni, la valeur par défaut est`EmrFSMetadata` .	Non
	La capacité de débit de lecture demandée pour la table de métadonnées. Si l'*READ_CAPA*	Non

Option	Description	Obligatoire
`-r READ_CAPACITY | --read-capacity READ_CAPACITY`	*CITY* argument n'est pas fourni, la valeur par défaut est400.	
`-w WRITE_CAPACITY | --write-capacity WRITE_CAPACITY`	La capacité de débit d'écriture demandée pour la table de métadonnées. Si l'*WRITE_CAPACITY* argument n'est pas fourni, la valeur par défaut est100.	Non

Example exemple d'emrfs set-metadata-capacity

L'exemple suivant définit la capacité de débit de lecture sur 600 et la capacité d'écriture sur 150 pour une table de métadonnées nommée EmrMetadataAlt.

```
$ emrfs set-metadata-capacity --metadata-name EmrMetadataAlt  --read-capacity 600 --
write-capacity 150
  read-capacity: 400
  write-capacity: 100
  status: UPDATING
  approximate-item-count (6 hour delay): 0
```

Sous commande delete-metadata emrfs

[Options] pour emrfs delete-metadata

Option	Description	Obligatoire
`-m METADATA_NAME | --metadata-name METADATA_NAME`	*METADATA_NAME* est le nom de la table de métadonnées DynamoDB. Si l'*METADATA_NAME* argument n'est pas fourni, la valeur par défaut estEmrFSMetadata .	Non

Example Exemple delete-metadata emrfs

L'exemple suivant supprime la table de métadonnées par défaut.

```
$ emrfs delete-metadata
```

Sous commande create-metadata emrfs

[Options] pour emrfs create-metadata

Option	Description	Obligatoire
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* est le nom de la table de métadonnées DynamoDB. Si l'*METADATA_NAME* argument n'est pas fourni, la valeur par défaut est`EmrFSMetadata` .	Non
`-r` *READ_CAPACITY* | `--read-capacity` *READ_CAPACITY*	La capacité de débit de lecture demandée pour la table de métadonnées. Si l'*READ_CAPACITY* argument n'est pas fourni, la valeur par défaut est`400`.	Non
`-w` *WRITE_CAPACITY* | `--write-capacity` *WRITE_CAPACITY*	La capacité de débit d'écriture demandée pour la table de métadonnées. Si l'*WRITE_CAPACITY* argument n'est pas fourni, la valeur par défaut est`100`.	Non

Example Exemple create-metadata emrfs

L'exemple suivant crée une table de métadonnée nommée `EmrFSMetadataAlt`.

```
$ emrfs create-metadata -m EmrFSMetadataAlt
Creating metadata: EmrFSMetadataAlt
EmrFSMetadataAlt
  read-capacity: 400
  write-capacity: 100
```

```
  status: ACTIVE
  approximate-item-count (6 hour delay): 0
```

sous-commande emrfs list-metadata-stores

La sous-commande emrfs list-metadata-stores n'a aucune [option].

Example List-metadata-stores exemple

L'exemple suivant répertorie vos tables de métadonnées.

```
$ emrfs list-metadata-stores
  EmrFSMetadata
```

Sous-commande diff emrfs

[Options] pour emrfs diff

Option	Description	Obligatoire
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* est le nom de la table de métadonnées DynamoDB. Si l'*METADATA_NAME* argument n'est pas fourni, la valeur par défaut est`EmrFSMetadata` .	Non
s3://s3Path	Le chemin d'accès au compartiment Amazon S3 à comparer à la table des métadonnées. Synchronisation de compartiments de façon récursive.	Oui

Example Exemple diff emrfs

L'exemple suivant compare la table de métadonnées par défaut dans un compartiment Amazon S3.

```
$ emrfs diff s3://elasticmapreduce/samples/cloudfront
BOTH | MANIFEST ONLY | S3 ONLY
DIR elasticmapreduce/samples/cloudfront
DIR elasticmapreduce/samples/cloudfront/code/
DIR elasticmapreduce/samples/cloudfront/input/
```

```
DIR elasticmapreduce/samples/cloudfront/logprocessor.jar
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-14.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-15.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-16.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-17.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-18.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-19.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-20.WxYz1234
DIR elasticmapreduce/samples/cloudfront/code/cloudfront-loganalyzer.tgz
```

Sous commande delete emrfs

[Options] pour emrfs delete

Option	Description	Obligatoire
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* est le nom de la table de métadonnées DynamoDB. Si l'*METADATA_NAME* argument n'est pas fourni, la valeur par défaut est`EmrFSMetadata` .	Non
s3://s3Path	Le chemin d'accès au compartiment Amazon S3 que vous suivez pour une vue cohérente. Synchronisation de compartiments de façon récursive.	Oui
`-t` *TIME* | `--time` *TIME*	L'heure d'expiration (interprétée à l'aide de l'argument d'unité de temps). Toutes les entrées de métadonnées antérieures à l'*TIME*argument sont supprimées pour le compartiment spécifié.	
`-u` *UNIT* | `--time-unit` *UNIT*	La mesure utilisée pour interpréter l'argument temps (nanosecondes, microsecondes, millisecondes, secondes, minutes, heures ou jours). Si aucun argument n'est spécifié, la valeur par défaut est `days`.	

Option	Description	Obligatoire
`--read-consumption` *READ_CONSUMPTION*	Le montant requis de débit de lecture disponible utilisé pour l'opération delete. Si l'*READ_CONSUMPTION* argument n'est pas spécifié, la valeur par défaut est`400`.	Non
`--write-consumption` *WRITE_CONSUMPTION*	Le montant requis de débit d'écriture disponible utilisé pour l'opération delete. Si l'*WRITE_CONSUMPTION* argument n'est pas spécifié, la valeur par défaut est`100`.	Non

Example Exemple delete emrfs

L'exemple suivant supprime tous les objets dans un compartiment Amazon S3 depuis les métadonnées de suivi pour une vue cohérente.

```
$ emrfs delete s3://elasticmapreduce/samples/cloudfront
entries deleted: 11
```

Sous commande import emrfs

[Options] pour emrfs import

Option	Description	Obligatoire
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* est le nom de la table de métadonnées DynamoDB. Si l'*METADATA_NAME* argument n'est pas fourni, la valeur par défaut est`EmrFSMetadata` .	Non
s3://s3Path	Le chemin d'accès au compartiment Amazon S3 que vous suivez pour une vue cohérente. Synchronisation de compartiments de façon récursive.	Oui

Option	Description	Obligatoire
`--read-consumption` *READ_CONSUMPTION*	Le montant requis de débit de lecture disponible utilisé pour l'opération delete. Si l'*READ_CONSUMPTION* argument n'est pas spécifié, la valeur par défaut est`400`.	Non
`--write-consumption` *WRITE_CONSUMPTION*	Le montant requis de débit d'écriture disponible utilisé pour l'opération delete. Si l'*WRITE_CONSUMPTION* argument n'est pas spécifié, la valeur par défaut est`100`.	Non

Example Exemple import emrfs

L'exemple suivant importe tous les objets dans un compartiment Amazon S3 avec les métadonnées de suivi pour une vue cohérente. Toutes les clés inconnus sont ignorées.

```
$ emrfs import s3://elasticmapreduce/samples/cloudfront
```

Sous-commande sync emrfs

[Options] pour emrfs sync

Option	Description	Obligatoire
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* est le nom de la table de métadonnées DynamoDB. Si l'*METADATA_NAME* argument n'est pas fourni, la valeur par défaut est`EmrFSMetadata` .	Non
s3://s3Path		Oui

Option	Description	Obligatoire
	Le chemin d'accès au compartiment Amazon S3 que vous suivez pour une vue cohérente. Synchronisation de compartiments de façon récursive.	
`--read-consumption` *READ_CONSUMPTION*	Le montant requis de débit de lecture disponible utilisé pour l'opération delete. Si l'*READ_CONSUMPTION* argument n'est pas spécifié, la valeur par défaut est`400`.	Non
`--write-consumption` *WRITE_CONSUMPTION*	Le montant requis de débit d'écriture disponible utilisé pour l'opération delete. Si l'*WRITE_CONSUMPTION* argument n'est pas spécifié, la valeur par défaut est`100`.	Non

Example Exemple commande sync emrfs

L'exemple suivant importe tous les objets dans un compartiment Amazon S3 avec les métadonnées de suivi pour une vue cohérente. Toutes les clés inconnues sont supprimées.

```
$ emrfs sync s3://elasticmapreduce/samples/cloudfront
Synching samples/cloudfront                                    0 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/code/                              1 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/                                   2 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/input/                             9 added | 0 updated |
 0 removed | 0 unchanged
Done synching s3://elasticmapreduce/samples/cloudfront         9 added | 0 updated |
 1 removed | 0 unchanged
creating 3 folder key(s)
folders written: 3
```

Sous-commande read-sqs emrfs

[Options] pour emrfs read-sqs

Option	Description	Obligatoire
`-q` *QUEUE_NAME* | `--queue-name` *QUEUE_NAME*	*QUEUE_NAME* est le nom de la file d'attente Amazon SQS configurée dans. `emrfs-site.xml` La valeur par défaut est **EMRFS-Inconsistency-<jobFlowId>** .	Oui
`-o` *OUTPUT_FILE* | `--output-file` *OUTPUT_FILE*	*OUTPUT_FILE* est le chemin d'accès au fichier de sortie sur le système de fichiers local du nœud maître. Les messages lus depuis la file d'attente sont écrits dans ce fichier.	Oui

Sous commande delete-sqs emrfs

[Options] pour emrfs delete-sqs

Option	Description	Obligatoire
`-q` *QUEUE_NAME* | `--queue-name` *QUEUE_NAME*	*QUEUE_NAME* est le nom de la file d'attente Amazon SQS configurée dans. `emrfs-site.xml` La valeur par défaut est **EMRFS-Inconsistency-<jobFlowId>** .	Oui

Soumission de commandes CLI EMRFS comme étapes

L'exemple suivant montre comment utiliser l'`emrfs`utilitaire sur le nœud principal en tirant parti de l'API AWS CLI or et en exécutant la `command-runner.jar` `emrfs` commande en tant qu'étape par étape. L'exemple utilise la commande AWS SDK pour Python (Boto3) pour ajouter une étape à un cluster qui ajoute des objets dans un compartiment Amazon S3 à la table de métadonnées EMRFS par défaut.

```
import boto3
from botocore.exceptions import ClientError


def add_emrfs_step(command, bucket_url, cluster_id, emr_client):
    """
    Add an EMRFS command as a job flow step to an existing cluster.

    :param command: The EMRFS command to run.
    :param bucket_url: The URL of a bucket that contains tracking metadata.
    :param cluster_id: The ID of the cluster to update.
    :param emr_client: The Boto3 Amazon EMR client object.
    :return: The ID of the added job flow step. Status can be tracked by calling
             the emr_client.describe_step() function.
    """
    job_flow_step = {
        "Name": "Example EMRFS Command Step",
        "ActionOnFailure": "CONTINUE",
        "HadoopJarStep": {
            "Jar": "command-runner.jar",
            "Args": ["/usr/bin/emrfs", command, bucket_url],
        },
    }

    try:
        response = emr_client.add_job_flow_steps(
            JobFlowId=cluster_id, Steps=[job_flow_step]
        )
        step_id = response["StepIds"][0]
        print(f"Added step {step_id} to cluster {cluster_id}.")
    except ClientError:
        print(f"Couldn't add a step to cluster {cluster_id}.")
        raise
    else:
        return step_id


def usage_demo():
    emr_client = boto3.client("emr")
    # Assumes the first waiting cluster has EMRFS enabled and has created metadata
    # with the default name of 'EmrFSMetadata'.
    cluster = emr_client.list_clusters(ClusterStates=["WAITING"])["Clusters"][0]
    add_emrfs_step(
```

```
        "sync", "s3://elasticmapreduce/samples/cloudfront", cluster["Id"], emr_client
    )


if __name__ == "__main__":
    usage_demo()
```

Vous pouvez utiliser la valeur `step_id` pour vérifier les journaux concernant le résultat de l'opération.

Autorisation d'accès aux données EMRFS dans Amazon S3

Par défaut, le rôle EMR pour EC2 détermine les autorisations d'accès aux données EMRFS dans Amazon S3. Les politique IAM qui sont attachées à ce rôle s'appliquent, quel que soit l'utilisateur ou le groupe qui effectuent la demande via le système EMRFS. L'argument par défaut est `EMR_EC2_DefaultRole`. Pour plus d'informations, consultez la section Rôle de service pour les EC2 instances de cluster (profil d'EC2 instance).

À partir de la version 5.10.0 d'Amazon EMR, vous pouvez utiliser une configuration de sécurité pour spécifier les rôles IAM pour EMRFS. Cela vous permet de personnaliser les autorisations pour les demandes EMRFS à Amazon S3 pour les clusters ayant plusieurs utilisateurs. Vous pouvez spécifier différents rôles IAM pour différents utilisateurs et groupes, et pour différents emplacements de compartiment Amazon S3 en fonction du préfixe dans Amazon S3. Quand EMRFS envoie une demande à Amazon S3 qui correspond à des utilisateurs, des groupes ou des emplacements que vous spécifiez, le cluster utilise le rôle correspondant que vous spécifiez au lieu du rôle EMR pour lequel vous spécifiez. EC2 Pour plus d'informations, consultez Configuration de rôles IAM pour les demandes EMRFS à Amazon S3.

Sinon, si les exigences de votre solution Amazon EMR ne sont pas couvertes par les rôles IAM pour EMRFS, vous pouvez définir une classe de fournisseur d'informations d'identification personnalisées, ce qui vous permet de personnaliser l'accès aux données EMRFS dans Amazon S3.

Création d'un fournisseur d'informations d'identification personnalisées pour les données EMRFS dans Amazon S3

Pour créer un fournisseur d'informations d'identification personnalisé, vous implémentez les classes AWSCredentialsProvider et Hadoop Configurable.

Pour obtenir une explication détaillée de l'approche, consultez Analyse sécurisée des données depuis un autre AWS compte avec EMRFS dans le blog AWS Big Data. Le billet de blog inclut un didacticiel qui parcourt le processusend-to-end, depuis la création des rôles IAM au lancement du cluster. Il fournit aussi un exemple de code Java qui implémente la classe fournisseur d'informations d'identification personnalisées.

La procédure de base est la suivante :

Pour spécifier un fournisseur d'informations d'identification personnalisées

1. Créez une classe fournisseur d'informations d'identification personnalisées compilée comme fichier JAR.
2. Exécutez un script comme action d'amorçage pour copier le fichier JAR du fournisseur d'informations d'identification personnalisées sur l'emplacement `/usr/share/aws/emr/emrfs/auxlib` du nœud principal du cluster. Pour plus d'informations sur les actions d'amorçage, consultez (Facultatif) Création d'actions d'amorçage pour installer des logiciels supplémentaires.
3. Personnalisez la classification `emrfs-site` pour spécifier la classe que vous implémentez dans le fichier JAR. Pour plus d'informations sur la spécification d'objets de configuration pour personnaliser les applications, consultez la section Configuration des applications dans le Guide de version Amazon EMR.

   L'exemple suivant illustre une commande `create-cluster` qui lance un cluster Hive avec les paramètres de configuration courants et qui inclut également :

   - Une action d'amorçage qui exécute le script, *`copy_jar_file.sh`*, enregistré sur *`amzn-s3-demo-bucket`* dans Amazon S3.
   - Une classification `emrfs-site` qui spécifie un fournisseur d'informations d'identification personnalisées défini dans le fichier JAR comme *`MyCustomCredentialsProvider`*

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --applications Name=Hive \
--bootstrap-actions '[{"Path":"s3://amzn-s3-demo-bucket/
copy_jar_file.sh","Name":"Custom action"}]' \
--ec2-attributes '{"KeyName":"MyKeyPair","InstanceProfile":"EMR_EC2_DefaultRole",\
"SubnetId":"subnet-xxxxxxxx","EmrManagedSlaveSecurityGroup":"sg-xxxxxxxx",\
"EmrManagedMasterSecurityGroup":"sg-xxxxxxxx"}' \
--service-role EMR_DefaultRole_V2 --enable-debugging --release-label emr-7.9.0 \
--log-uri 's3n://amzn-s3-demo-bucket/' --name 'test-awscredentialsprovider-emrfs' \
--instance-type=m5.xlarge --instance-count 3  \
--configurations '[{"Classification":"emrfs-site",\
"Properties":
{"fs.s3.customAWSCredentialsProvider":"MyAWSCredentialsProviderWithUri"},\
"Configurations":[]}]'
```

Gestion du point de AWS Security Token Service terminaison par défaut

EMRFS utilise AWS Security Token Service (STS) pour récupérer des informations d'identification de sécurité temporaires afin d'accéder à vos AWS ressources. Les versions antérieures d'Amazon EMR envoient toutes les AWS STS demandes à un seul point de terminaison mondial à l'adresse. `https://sts.amazonaws.com` Amazon EMR publie les versions 5.31.0 et 6.1.0 et ultérieures pour envoyer des demandes aux points de terminaison régionaux. AWS STS Cela réduit la latence et améliore la validité des jetons de session. Pour plus d'informations sur les AWS STS points de terminaison, consultez la section Gestion AWS STS dans une AWS région du Guide de l'AWS Identity and Access Management utilisateur.

Lorsque vous utilisez les versions 5.31.0 et 6.1.0 et ultérieures d'Amazon EMR, vous pouvez remplacer le point de terminaison par défaut AWS STS . Pour ce faire, vous devez modifier la propriété `fs.s3.sts.endpoint` dans votre configuration `emrfs-site`.

L' AWS CLI exemple suivant définit le point de AWS STS terminaison par défaut utilisé par EMRFS comme point de terminaison global.

```
aws emr create-cluster --release-label <emr-5.33.0> --instance-type m5.xlarge \
--emrfs Args=[fs.s3.sts.endpoint=https://sts.amazonaws.com]
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

Vous pouvez également créer un fichier de configuration JSON à l'aide de l'exemple suivant et le spécifier à l'aide de l'argument `--configurations` de `emr create-cluster`. Pour plus d'informations sur l'utilisation de `--configurations`, consultez la Référence des commandes AWS CLI .

```
[
  {
    "classification": "emrfs-site",
    "properties": {
      "fs.s3.sts.endpoint": "https://sts.amazonaws.com"
    }
  }
]
```

Spécification du chiffrement Amazon S3 en utilisant les propriétés EMRFS

> **Important**
>
> À partir de la version 4.8.0 d'Amazon EMR, vous pouvez utiliser les configurations de sécurité pour appliquer les paramètres de chiffrement plus facilement et avec plus d'options. Nous vous recommandons d'utiliser les configurations de sécurité. Pour plus d'informations, consultez Configuration du chiffrement des données. Les instructions de console décrites dans cette section sont disponibles pour les versions antérieures à la version 4.8.0. Si vous utilisez l' AWS CLI pour configurer le chiffrement Amazon S3 à la fois dans la configuration de cluster et dans une configuration de sécurité dans des versions ultérieures, la configuration de sécurité prévaut.

Lorsque vous créez un cluster, vous pouvez spécifier un chiffrement côté serveur (SSE) ou chiffrement côté client (CSE) pour les données EMRFS dans Amazon S3 à l'aide de la console `emrfs-site` ou des propriétés de classification par le biais de l'ou du kit SDK EMR. AWS CLI Le chiffrement SSE et le chiffrement CSE pour Amazon S3 s'excluent mutuellement. Vous pouvez choisir l'un ou l'autre, mais pas les deux.

Pour AWS CLI obtenir des instructions, consultez la section correspondant à votre type de chiffrement ci-dessous.

Pour spécifier les options de chiffrement EMRFS à l'aide de l' AWS Management Console

1. Accédez à la nouvelle console Amazon EMR et sélectionnez Changer pour l'ancienne console depuis le menu latéral. Pour plus d'informations sur ce qu'implique le passage à l'ancienne console, consultez la rubrique Utilisation de l'ancienne console.
2. Choisissez Créer un cluster et Go to advanced options (Aller aux options avancées).
3. Choisissez une Release (Version) 4.7.2 ou antérieure.
4. Choisissez d'autres options pour Software and Steps (Logiciel et étapes) selon les besoins de votre application, puis cliquez sur Next (Suivant).
5. Choisissez les paramètres dans les volets Hardware (Matériel) et General Cluster Settings (Paramètres de cluster généraux) selon les besoins de votre application.
6. Dans le volet Security (Sécurité), sous Authentication and encryption (Authentification et chiffrement), sélectionnez l'option S3 Encryption (with EMRFS) (Chiffrement S3 (avec EMRFS)) à utiliser.

   > **Note**
   >
   > S3 server-side encryption with KMS Key Management (Chiffrement côté serveur S3 avec gestion de clé KMS) (SSE-KMS) n'est pas disponible lorsque vous utilisez Amazon EMR version 4.4 ou antérieure.

   - Si vous choisissez une option qui utilise AWS Key Management, choisissez un AWS KMS Key ID. Pour de plus amples informations, veuillez consulter Utilisation de AWS KMS keys pour le chiffrement EMRFS.
   - Si vous choisissez le S3 client-side encryption with custom materials provider (Chiffrement côté client S3 avec fournisseur de matériaux personnalisé), spécifiez le Class name (Nom de

classe) et l'JAR location (emplacement du JAR). Pour de plus amples informations, veuillez consulter Chiffrement côté client sur Amazon S3.

7. Choisissez d'autres options selon les besoins de votre application, puis choisissez Create Cluster (Créer un cluster).

Utilisation de AWS KMS keys pour le chiffrement EMRFS

La clé de AWS KMS chiffrement doit être créée dans la même région que votre instance de cluster Amazon EMR et que les compartiments Amazon S3 utilisés avec EMRFS. Si la clé que vous spécifiez se trouve dans un compte différent de celui que vous utilisez pour configurer un cluster, vous devez spécifier la clé à l'aide de son ARN.

Le rôle du profil d' EC2 instance Amazon doit être autorisé à utiliser la clé KMS que vous spécifiez. Le rôle par défaut du profil d'instance dans Amazon EMR est `EMR_EC2_DefaultRole`. Si vous utilisez un rôle différent pour le profil d'instance, ou si vous utilisez des rôles IAM pour les demandes EMRFS adressées à Amazon S3, assurez-vous que chaque rôle est ajouté en tant qu'utilisateur clé, le cas échéant. Cela donne au rôle des autorisations pour utiliser la clé KMS. Pour plus d'informations, consultez les sections Utilisation des stratégies de clé dans le Guide du développeur AWS Key Management Service et Configuration des rôles IAM pour les demandes EMRFS adressées à Amazon S3.

Vous pouvez utiliser AWS Management Console pour ajouter votre profil d'instance ou votre profil d' EC2 instance à la liste des utilisateurs clés pour la clé KMS spécifiée, ou vous pouvez utiliser l' AWS CLI ou un AWS kit SDK pour associer une stratégie de clé appropriée.

Notez qu'Amazon EMR prend uniquement en charge les clés KMS symétriques. Vous ne pouvez pas utiliser une clé KMS asymétrique pour chiffrer les données au repos dans un cluster Amazon EMR. Pour savoir si une clés KMS est symétrique ou asymétrique, consultez Identification de clés KMS symétriques et asymétriques.

La procédure ci-dessous décrit comment ajouter le profil d'instance Amazon EMR par défaut, `EMR_EC2_DefaultRole`, en tant qu'utilisateur de clé à l'aide de la AWS Management Console. Elle suppose que vous avez déjà créé une clé KMS. Pour créer une nouvelle clé KMS, consultez Création de clés dans le Guide du développeur AWS Key Management Service .

Pour ajouter le profil d' EC2 instance pour Amazon EMR à la liste des utilisateurs de clés de chiffrement

1. Connectez-vous à la console AWS Key Management Service (AWS KMS) AWS Management Console et ouvrez-la à l'adresse https://console.aws.amazon.com/kms.
2. Pour changer de Région AWS, utilisez le sélecteur de région dans l'angle supérieur droit de la page.
3. Sélectionnez l'alias de la clé KMS à modifier.
4. Sur la page de détails de la clé, sous Key Users (Utilisateurs de clés), choisissez Add (Ajouter).
5. Dans la boîte de dialogue Ajouter des utilisateurs clés sélectionnez le rôle approprié. Le nom du rôle par défaut est `EMR_EC2_DefaultRole`.
6. Choisissez Ajouter.

Chiffrement côté serveur sur Amazon S3

Lorsque vous configurez le chiffrement côté serveur sur Amazon S3, Amazon S3 chiffre les données au niveau de l'objet au moment où elles sont écrites sur le disque et déchiffre les données lorsqu'elles sont accédées. Pour plus d'informations sur SSE, consultez Protection des données à l'aide du chiffrement côté serveur dans le Guide de l'utilisateur Amazon Simple Storage Service.

Lorsque vous indiquez le chiffrement SSE sur Amazon EMR, vous pouvez choisir entre deux systèmes de gestion de clés différents :

- SSE-S3 – Amazon S3 gère les clés pour vous.
- SSE-KMS — Vous utilisez une clé AWS KMS key pour configurer des politiques adaptées à Amazon EMR. Pour plus d'informations sur les exigences clés pour Amazon EMR, consultez la section Utilisation des à des AWS KMS keys fins de chiffrement.

Le chiffrement SSE avec des clés fournies par le client (SSE-C) n'est pas disponible pour une utilisation avec Amazon EMR.

Pour créer un cluster avec SSE-S3 activé à l'aide du AWS CLI

- Saisissez la commande suivante :

```
aws emr create-cluster --release-label emr-4.7.2 or earlier \
```

```
--instance-count 3 --instance-type m5.xlarge --emrfs Encryption=ServerSide
```

Vous pouvez également activer SSE-S3 en définissant le fs.s3. enableServerSideLa propriété de chiffrement prend la valeur true dans les `emrfs-site` propriétés. Voir l'exemple pour SSE-KMS ci-dessous et omettez la propriété pour l'ID de clé.

Pour créer un cluster avec SSE-KMS activé à l'aide du AWS CLI

> **Note**
>
> SSE-KMS est uniquement disponible dans Amazon EMR version 4.5.0 et versions ultérieures.

- Tapez la AWS CLI commande suivante pour créer un cluster avec SSE-KMS, où se *keyID* trouve, par exemple AWS KMS key, un : *a4567b8-9900-12ab-1234-123a45678901*

  ```
  aws emr create-cluster --release-label emr-4.7.2 or earlier --instance-count 3 \
  --instance-type m5.xlarge --use-default-roles \
  --emrfs Encryption=ServerSide,Args=[fs.s3.serverSideEncryption.kms.keyId=keyId]
  ```

  --OR--

  Tapez la AWS CLI commande suivante à l'aide de la `emrfs-site` classification et fournissez un fichier de configuration JSON dont le contenu est similaire `myConfig.json` à celui de l'exemple ci-dessous :

  ```
  aws emr create-cluster --release-label emr-4.7.2 or earlier --instance-count 3
   --instance-type m5.xlarge --applications Name=Hadoop --configurations file://
  myConfig.json --use-default-roles
  ```

  Exemple de contenu de myConfig.json :

  ```
  [
    {
      "Classification":"emrfs-site",
      "Properties": {
         "fs.s3.enableServerSideEncryption": "true",
         "fs.s3.serverSideEncryption.kms.keyId":"a4567b8-9900-12ab-1234-123a45678901"
  ```

```
    }
  }
]
```

Propriétés de configuration pour SSE-S3 et SSE-KMS

Ces propriétés peuvent être configurées à l'aide de la classification de configuration `emrfs-site`. SSE-KMS est uniquement disponible dans Amazon EMR version 4.5.0 et versions ultérieures.

Propriété	Valeur par défaut	Description
`fs.s3.enableServerSideEncryption`	**`false`**	Lorsque la valeur est **`true`**, les objets stockés dans Amazon S3 sont chiffrés à l'aide du chiffrement côté serveur. Si aucune clé n'est spécifiée, SSE-S3 est utilisé.
`fs.s3.serverSideEncryption.kms.keyId`	**`n/a`**	Spécifie un ID de AWS KMS clé ou un ARN. Si une clé est spécifiée, SSE-KMS est utilisé.

Chiffrement côté client sur Amazon S3

Avec le chiffrement côté client sur Amazon S3, le chiffrement et le déchiffrement par Amazon S3 se déroulent dans le client EMRFS de votre cluster. Les objets sont chiffrés avant d'être chargés sur Amazon S3 et déchiffrés après leur chargement. Le fournisseur que vous indiquez fournit la clé de chiffrement utilisée par le client. Le client peut utiliser les clés fournies par AWS KMS (CSE-KMS) ou une classe Java personnalisée qui fournit la clé racine côté client (CSE-C). Les spécificités du chiffrement sont légèrement différentes entre CSE-KMS et CSE-C, en fonction du fournisseur indiqué et des métadonnées de l'objet à déchiffrer ou à chiffrer. Pour plus d'informations sur ces différences, consultez Protection des données à l'aide du chiffrement côté client dans le Guide de l'utilisateur Amazon Simple Storage Service.

> **Note**
>
> Le chiffrement CSE sur Amazon S3 garantit uniquement que les données EMRFS échangées avec Amazon S3 sont chiffrées ; cela ne signifie pas que toutes les données sur les volumes des instances du cluster sont chiffrées. De plus, étant donné que Hue n'utilise pas EMRFS, les objets que le navigateur de fichiers S3 de Hue écrit sur Amazon S3 ne sont pas chiffrés.

Pour spécifier CSE-KMS pour les données EMRFS dans Amazon S3 à l'aide de l' AWS CLI

- Tapez la commande suivante et *MyKMSKeyID* remplacez-la par l'ID ou l'ARN de la clé KMS à utiliser :

```
aws emr create-cluster --release-label emr-4.7.2 or earlier
--emrfs Encryption=ClientSide,ProviderType=KMS,KMSKeyId=MyKMSKeyId
```

Création d'un fournisseur de clés personnalisé

Selon le type de chiffrement que vous utilisez lors de la création d'un fournisseur de clés personnalisé, l'application doit également implémenter différentes EncryptionMaterialsProvider interfaces. Les deux interfaces sont disponibles dans le AWS SDK pour Java version 1.11.0 et versions ultérieures.

- Pour implémenter le chiffrement Amazon S3, utilisez le fichier com.amazonaws.services.s3.model. EncryptionMaterialsProvider interface.
- Pour implémenter le chiffrement du disque local, utilisez le fichier com.amazonaws.services.elasticmapreduce.spi.security. EncryptionMaterialsProvider interface.

Vous pouvez utiliser n'importe quelle stratégie pour fournir des matériaux de chiffrement pour la mise en œuvre. Par exemple, vous pouvez choisir de fournir des matériaux de chiffrement statiques ou d'intégrer un système de gestion des clés plus complexe.

Si vous utilisez le chiffrement Amazon S3, vous devez utiliser les algorithmes de chiffrement AES/GCM/NoPaddingpour le matériel de chiffrement personnalisé.

Si vous utilisez le chiffrement de disque local, l'algorithme de chiffrement à utiliser pour le matériel de chiffrement personnalisé varie selon la version de l'EMR. Pour Amazon EMR 7.0.0 et versions antérieures, vous devez utiliser. AES/GCM/NoPadding Pour Amazon EMR 7.1.0 et versions ultérieures, vous devez utiliser AES.

La EncryptionMaterialsProvider classe obtient les matériaux de chiffrement par contexte de chiffrement. Amazon EMR renseigne les informations contextuelles de chiffrement au moment de l'exécution pour aider l'appelant à déterminer les matériaux de chiffrement à renvoyer.

Example Exemple : utilisation d'un fournisseur de clés personnalisé pour le chiffrement Amazon S3 avec EMRFS

Lorsqu'Amazon EMR extrait les matériaux de chiffrement de la EncryptionMaterialsProvider classe pour effectuer le chiffrement, EMRFS remplit éventuellement l'argument materialsDescription avec deux champs : l'URI Amazon S3 de l'objet et le cluster, qui peuvent être utilisés par la classe pour JobFlowId renvoyer les matériaux de chiffrement de manière sélective. EncryptionMaterialsProvider

Par exemple, le fournisseur peut renvoyer des clés différentes pour différents préfixes d'URI Amazon S3. C'est la description des matériaux de chiffrement renvoyés qui est finalement stockée avec l'objet Amazon S3 plutôt que la valeur materialsDescription qui est générée par EMRFS et transmise au fournisseur. Lors du déchiffrement d'un objet Amazon S3, la description des matériaux de chiffrement est transmise à la EncryptionMaterialsProvider classe, afin qu'elle puisse, à nouveau, renvoyer sélectivement la clé correspondante pour déchiffrer l'objet.

Une implémentation EncryptionMaterialsProvider de référence est fournie ci-dessous. Un autre fournisseur personnalisé est disponible auprès de GitHub. EMRFSRSAEncryptionMaterialsProvider

```
import com.amazonaws.services.s3.model.EncryptionMaterials;
import com.amazonaws.services.s3.model.EncryptionMaterialsProvider;
import com.amazonaws.services.s3.model.KMSEncryptionMaterials;
import org.apache.hadoop.conf.Configurable;
import org.apache.hadoop.conf.Configuration;

import java.util.Map;

/**
 * Provides KMSEncryptionMaterials according to Configuration
 */
public class MyEncryptionMaterialsProviders implements EncryptionMaterialsProvider,
 Configurable{
  private Configuration conf;
```

```
  private String kmsKeyId;
  private EncryptionMaterials encryptionMaterials;

  private void init() {
    this.kmsKeyId = conf.get("my.kms.key.id");
    this.encryptionMaterials = new KMSEncryptionMaterials(kmsKeyId);
  }

  @Override
  public void setConf(Configuration conf) {
    this.conf = conf;
    init();
  }

  @Override
  public Configuration getConf() {
    return this.conf;
  }

  @Override
  public void refresh() {

  }

  @Override
  public EncryptionMaterials getEncryptionMaterials(Map<String, String>
 materialsDescription) {
    return this.encryptionMaterials;
  }

  @Override
  public EncryptionMaterials getEncryptionMaterials() {
    return this.encryptionMaterials;
  }
}
```

Spécification d'un fournisseur de matériaux personnalisés à l'aide de l' AWS CLI

Pour utiliser l' AWS CLI, transmettez les arguments `Encryption`, `ProviderType`, `CustomProviderClass` et `CustomProviderLocation` à l'option `emrfs`.

```
aws emr create-cluster --instance-type m5.xlarge --release-label emr-4.7.2 or earlier
 --emrfs Encryption=ClientSide,ProviderType=Custom,CustomProviderLocation=s3://amzn-s3-
demo-bucket/myfolder/provider.jar,CustomProviderClass=classname
```

L'affectation à `ClientSide` de la valeur `Encryption` rend possible le chiffrement côté client. `CustomProviderClass` est le nom de votre objet `EncryptionMaterialsProvider` et `CustomProviderLocation` est l'emplacement local ou Amazon S3 à partir duquel Amazon EMR copie `CustomProviderClass` sur chaque nœud du cluster et le place dans le chemin de classe.

Spécification d'un fournisseur de matériaux personnalisés à l'aide d'un kit SDK

Pour utiliser un kit SDK, vous pouvez définir la propriété `fs.s3.cse.encryptionMaterialsProvider.uri` pour télécharger la classe `EncryptionMaterialsProvider` personnalisée que vous stockez dans Amazon S3 vers chaque nœud de votre cluster. Vous configurez cela dans le fichier `emrfs-site.xml` avec CSE activé et l'emplacement correct du fournisseur personnalisé.

Par exemple, lors de l' AWS SDK pour Java utilisation RunJobFlowRequest, votre code peut ressembler à ce qui suit :

```
<snip>
  Map<String,String> emrfsProperties = new HashMap<String,String>();
      emrfsProperties.put("fs.s3.cse.encryptionMaterialsProvider.uri","s3://amzn-s3-
demo-bucket/MyCustomEncryptionMaterialsProvider.jar");
      emrfsProperties.put("fs.s3.cse.enabled","true");
      emrfsProperties.put("fs.s3.consistent","true");

 emrfsProperties.put("fs.s3.cse.encryptionMaterialsProvider","full.class.name.of.EncryptionMate

  Configuration myEmrfsConfig = new Configuration()
      .withClassification("emrfs-site")
      .withProperties(emrfsProperties);

  RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("Custom EncryptionMaterialsProvider")
   .withReleaseLabel("emr-7.9.0")
   .withApplications(myApp)
   .withConfigurations(myEmrfsConfig)
   .withServiceRole("EMR_DefaultRole_V2")
   .withJobFlowRole("EMR_EC2_DefaultRole")
   .withLogUri("s3://myLogUri/")
```

```
    .withInstances(new JobFlowInstancesConfig()
     .withEc2KeyName("myEc2Key")
     .withInstanceCount(2)
     .withKeepJobFlowAliveWhenNoSteps(true)
     .withMasterInstanceType("m5.xlarge")
     .withSlaveInstanceType("m5.xlarge")
    );

  RunJobFlowResult result = emr.runJobFlow(request);
</snip>
```

Personnalisé EncryptionMaterialsProvider avec arguments

Vous devrez peut-être transmettre des arguments directement au fournisseur. Pour ce faire, vous pouvez utiliser la classification de configuration `emrfs-site` avec les propriétés définies comme arguments personnalisés. Un exemple de configuration est indiqué ci-dessous, qui est enregistré dans un fichier, `myConfig.json` :

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "myProvider.arg1":"value1",
     "myProvider.arg2":"value2"
      }
    }
 ]
```

À l'aide de la `create-cluster` commande du AWS CLI, vous pouvez utiliser l'`--configurations`option pour spécifier le fichier comme indiqué ci-dessous :

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge
 --instance-count 2 --configurations file://myConfig.json --emrfs
 Encryption=ClientSide,CustomProviderLocation=s3://amzn-s3-demo-bucket/myfolder/
myprovider.jar,CustomProviderClass=classname
```

Configuration de la prise en charge d'EMRFS S3EC V2

Les versions du SDK Java S3 (1.11.837 et versions ultérieures) prennent en charge la version 2 du client de chiffrement (S3EC V2) avec diverses améliorations de sécurité. Pour plus d'informations,

consultez l'article de blog S3 Updates to the Amazon S3 encryption client. Reportez-vous également à Migration du client de chiffrement Amazon S3 dans le Guide du AWS SDK pour Java développeur.

Le client de chiffrement V1 est toujours disponible dans le SDK pour des raisons de rétrocompatibilité. Par défaut, EMRFS utilise S3EC V1 pour chiffrer et déchiffrer les objets S3 si le CSE est activé.

Les objets S3 chiffrés avec S3EC V2 ne peuvent pas être déchiffrés par EMRFS sur un cluster EMR dont la version de publication est antérieure à emr-5.31.0 (emr-5.30.1 et versions antérieures, emr-6.1.0 et versions antérieures).

Example Configurer EMRFS pour utiliser S3EC V2

Pour configurer EMRFS afin d'utiliser S3EC V2, ajoutez la configuration suivante :

```
{
  "Classification": "emrfs-site",
  "Properties": {
    "fs.s3.cse.encryptionV2.enabled": "true"
  }
}
```

Propriétés **`emrfs-site.xml`** pour le chiffrement côté client Amazon S3

Propriété	Valeur par défaut	Description
`fs.s3.cse.enabled`	**`false`**	Lorsque la valeur est **`true`**, les objets EMRFS stockés dans Amazon S3 sont chiffrés à l'aide du chiffrement côté client.
`fs.s3.cse.encryptionV2.enabled`	**`false`**	Lorsqu'il est défini sur `true`, EMRFS utilise le client de chiffrement S3 version 2 pour chiffrer et déchiffrer les objets sur S3. Disponible à partir de la version 5.31.0 d'EMR.

Propriété	Valeur par défaut	Description
`fs.s3.cse.encryptionMaterialsProvider.uri`	**N/A**	S'applique lors de l'utilisation de supports de chiffrement personnalisés. L'URI Amazon S3 où se trouve le fichier JAR contenant le `EncryptionMaterialsProvider` . Lorsque vous fournissez cet URI, Amazon EMR télécharge automatiquement le fichier JAR sur tous les nœuds du cluster.
`fs.s3.cse.encryptionMaterialsProvider`	**N/A**	Le chemin de classe `EncryptionMaterialsProvider` utilisé avec le chiffrement côté client. Lorsque vous utilisez CSE-KMS, spécifiez `com.amazon.ws.emr.hadoop.fs.cse.KMSEncryptionMaterialsProvider` .
`fs.s3.cse.materialsDescription.enabled`	**false**	Lorsque la valeur est définie sur`true`, elle remplit l'attribut materialsDescription des objets chiffrés avec l'URI Amazon S3 de l'objet et le. JobFlowId Définissez sur `true` lors de l'utilisation de supports de chiffrement personnalisés.

Propriété	Valeur par défaut	Description
`fs.s3.cse.kms.keyId`	**N/A**	S'applique lors de l'utilisation de CSE-KMS. La valeur de KeyId, ARN, ou l'alias de la clé KMS utilisée pour le chiffrement.
`fs.s3.cse.cryptoStorageMode`	**ObjectMetadata**	Le mode de stockage Amazon S3. Par défaut, la description des informations de chiffrement est stockée dans les métadonnées de l'objet. Vous pouvez également stocker la description dans un fichier d'instructions. Les valeurs valides sont ObjectMetadata etInstructionFile. Pour de plus d'informations, consultez Chiffrement des données côté client avec AWS SDK pour Java et Amazon S3.

Chiffrement côté client Amazon S3 avec S3A

À partir de la version 7.6.0 d'Amazon Elastic Map Reduce (EMR), le connecteur de système de fichiers S3A prend désormais en charge le chiffrement côté client Amazon S3. Cela signifie que le chiffrement et le déchiffrement des données Amazon S3 se font directement dans le client S3A de votre cluster informatique. Lorsque vous utilisez cette fonctionnalité, les fichiers sont automatiquement chiffrés avant d'être chargés sur Amazon S3 et déchiffrés lors du téléchargement. Pour obtenir des informations complètes sur la méthodologie de chiffrement et sa mise en œuvre, les utilisateurs peuvent se référer à la section Protection des données à l'aide du chiffrement côté client dans le guide de l'utilisateur d'Amazon Simple Storage Service.

Lorsque vous activez le chiffrement côté client (CSE) avec S3A dans Amazon EMR, vous disposez de deux options de système de gestion des clés :

- CSE-KMS — Cette approche utilise une clé AWS Key Management Service (KMS) configurée avec des politiques spécialement conçues pour Amazon EMR. Pour obtenir des informations détaillées sur les exigences relatives aux clés, reportez-vous à la documentation sur l'utilisation des AWS KMS clés pour le chiffrement.
- CSE-CUSTOM — Cette méthode vous permet d'intégrer une classe Java personnalisée qui fournit la clé racine côté client chargée du chiffrement et du déchiffrement des données.

> **Note**
>
> Le chiffrement côté client S3A dans EMR est intrinsèquement compatible avec le chiffrement côté client EMRFS, ce qui signifie que les objets chiffrés à l'aide d'EMRFS CSE peuvent être lus via S3A CSE.

Rubriques

- Configuration de CSE-KMS
- Configuration de CSE-CUSTOM
- Propriétés du chiffrement côté client Amazon S3 avec S3A

Configuration de CSE-KMS

Vous pouvez activer le chiffrement côté client à l'aide de AWS KMS (CSE-KMS) dans deux domaines principaux :

- Le premier est la configuration à l'échelle du cluster :

```
[
  {
    "Classification":"core-site",
    "Properties": {
       "fs.s3a.encryption.algorithm": "CSE-KMS",
       "fs.s3a.encryption.key":"${KMS_KEY_ID}",
    }
  }
]
```

> **Note**
>
> Si la région AWS KMS clé est différente de la région du compartiment S3/EMR, vous devez définir la configuration supplémentaire suivante :. `fs.s3a.encryption.cse.kms.region=${KMS_REGION}`

- La seconde est la configuration spécifique à la tâche ou à l'application. Le CSE-KMS peut être configuré pour une application Spark spécifique comme suit :

```
spark-submit --conf spark.hadoop.fs.s3a.encryption.algorithm=CSE-KMS --conf
 spark.hadoop.fs.s3a.encryption.key=${KMS_KEY_ID}
```

Configuration de CSE-CUSTOM

Pour utiliser CSE-CUSTOM, vous devez créer un fournisseur de clés personnalisé en implémentant l'interface Keyring. Voici un exemple de mise en œuvre :

```
public class CustomKeyring implements Keyring {
  public CustomKeyring()  {
    // custom code
  }
```

```
  @Override
  public EncryptionMaterials onEncrypt(EncryptionMaterials encryptionMaterials) {
    // custom code
  }

  @Override
  public DecryptionMaterials onDecrypt(DecryptionMaterials decryptionMaterials,
      ListEncryptedDataKey list) {
    // custom code
  }
```

Vous pouvez activer les clés personnalisées de chiffrement côté client (CSE-CUSTOM) dans deux domaines principaux :

- Le premier champ d'application est la configuration à l'échelle du cluster :

```
[
  {
    "Classification":"core-site",
    "Properties": {
       "fs.s3a.encryption.algorithm": "CSE-CUSTOM",
       "fs.s3a.cse.customKeyringProvider.uri":"S3 path of custom jar",
       "fs.s3a.encryption.cse.custom.keyring.class.name":"fully qualified class name"
    }
  }
]
```

- La seconde est la configuration spécifique à la tâche ou à l'application. CSE-CUSTOM peut être configuré pour une application Spark spécifique comme suit :

```
spark-submit --conf spark.hadoop.fs.s3a.encryption.algorithm=CSE-CUSTOM --conf
 spark.hadoop.fs.s3a.encryption.cse.custom.keyring.class.name=fully qualified class
 name
```

> **Note**
>
> Assurez-vous que le fichier JAR personnalisé requis pour générer les clés de chiffrement/déchiffrement est présent dans le chemin de classe.

Propriétés du chiffrement côté client Amazon S3 avec S3A

Pour configurer le chiffrement côté client avec S3A, plusieurs propriétés de configuration doivent être définies dans les paramètres de votre fichier core-site.xml. Pour plus d'informations sur les paramètres de configuration personnalisés, voir Configuration des applications.

Propriété	Valeur par défaut	Description
fs.s3a.encryption.algorithm	N/A	Lorsqu'ils sont définis sur CSE-KMS ou CSE-CUSTOM , les objets stockés dans Amazon S3 sont chiffrés à l'aide du chiffrement côté client.
fs.s3a.encryption.key	N/A	S'applique lors de l'utilisation CSE-KMS. La valeur de l' KeyIdARN ou de l'alias de la clé KMS utilisée pour le chiffrement.
fs.s3a.encryption.cse.kms.region	N/A	S'applique lors de l'utilisation CSE-KMS. Région dans laquelle la AWS KMS clé est générée. Par défaut, la région KMS est définie sur des valeurs identiques à celles de la région du compartiment S3 ou du cluster EMR.
fs.s3a.encryption.cse.custom.keyring.class.name	N/A	S'applique lors de l'utilisation CSE-KMS. Le nom de classe complet du fournisseur de clés personnalisé.
fs.s3a.cse. customKeyringProvider.uri	N/A	S'applique lors de l'utilisation CSE-CUSTOM . L'URI Amazon S3 où se trouve

Propriété	Valeur par défaut	Description
		le fichier JAR contenant l'implémentation personnal isée de Keyring. Lorsque vous fournissez cet URI, Amazon EMR télécharge automatiq uement le fichier JAR sur tous les nœuds du cluster.
fs.s3a.encryption.cse.v1.co mpatibility.enabled	`'true'`	Cela permet une rétrocomp atibilité avec les anciens clients du SDK, comme celui utilisé avec EMRFS. Désactive z-la lorsque cette dépendance n'existe pas, pour de meilleure s performances.

CloudWatch Agent Amazon

L' CloudWatch agent Amazon sur Amazon EMR est un outil qui permet de surveiller les EC2 instances Amazon de votre cluster EMR. Vous pouvez stocker et consulter les métriques que vous collectez avec l' CloudWatch agent CloudWatch. Pour plus d'informations sur l' CloudWatch agent, consultez le guide de CloudWatch l'utilisateur Amazon.

> **Note**
>
> Des frais supplémentaires vous seront facturés si vous en utilisez un autre Services AWS pour publier, consulter ou consulter les statistiques des CloudWatch agents Amazon. Consultez les pages suivantes pour plus d'informations sur les tarifs.
>
> - CloudWatch Tarifs Amazon
> - Amazon Managed Service for Prometheus
> - Amazon Managed Grafana
>
> Ces métriques sont distinctes des métriques gratuites qu'Amazon EMR fournit sous l'espace de `AWS/ElasticMapReduce` noms dans. CloudWatch Pour plus d'informations sur les métriques que l' CloudWatch agent ne publie pas, consultez la section Surveillance des métriques Amazon EMR avec. CloudWatch

Avec Amazon EMR 7.0 et versions ultérieures, vous pouvez installer une version personnalisée de l' CloudWatch agent Amazon sur votre cluster pour collecter des métriques à partir de votre cluster EMR. CloudWatchAgent est pris en charge sur les clusters de rôles d'exécution pour EMR 7.6 et versions ultérieures. L'agent collecte les métriques de tous les nœuds du cluster, les rassemble sur le nœud primaire et publie les métriques de ce nœud dans le cloud.

Vous pouvez installer l'agent lorsque vous créez un cluster dans la console ou lorsque vous utilisez l'API `create-cluster`. Pour de plus amples informations, veuillez consulter Créez un cluster EMR qui utilise l'agent Amazon Amazon Amazon CloudWatch . Par défaut, un cluster Amazon EMR qui exécute la série 7.x publie 34 métriques CloudWatch au niveau du système toutes les 60 secondes, mais vous pouvez configurer l'agent pour qu'il publie différentes métriques. Une autre option consiste à publier des métriques sur Amazon Managed Service for Prometheus, et vous pouvez également choisir les métriques à publier. Selon les cas d'utilisation et les configurations, vous pouvez configurer

l'agent pour consulter et les interroger dans la CloudWatch console, Amazon Managed Grafana, ou via APIs la for CloudWatch ou Amazon Managed Service for Prometheus. Ils Services AWS entraînent des frais lorsque vous les utilisez pour stocker et interroger des métriques.

L' CloudWatch agent sur Amazon EMR peut publier des métriques système ainsi que des métriques JMX pour les services suivants sur votre cluster Amazon EMR.

- Hadoop DataNode
- Hadoop NameNode
- Laine NodeManager
- Laine ResourceManager
- HBase Master — Amazon EMR 7.1 ou version ultérieure uniquement
- HBase RegionServer — Amazon EMR 7.1 et versions ultérieures uniquement
- HBase ThriftServer — Amazon EMR 7.1 et versions ultérieures uniquement

Pour plus d'informations sur les métriques disponibles et sur la façon de configurer l' CloudWatch agent sur Amazon EMR, consultez Configuration de l' CloudWatch agent pour Amazon EMR.

Le tableau suivant répertorie la version AmazonCloudWatchAgent incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants avec lesquels Amazon EMR installe. AmazonCloudWatchAgent

Pour connaître la version des composants installés avec AmazonCloudWatchAgent cette version, consultez les versions des composants de la version 7.9.0.

AmazonCloudWatchAgent informations de version pour emr-7.9.0

Étiquette de version Amazon EMR	AmazonCloudWatchAgent Version	Composants installés avec AmazonCloudWatchAgent
emr-7.9.0	AmazonCloudWatchAgent 1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno

Étiquette de version Amazon EMR	AmazonCloudWatchAgent Version	Composants installés avec AmazonCloudWatchAgent
		de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Rubriques

- Créez un cluster EMR qui utilise l'agent Amazon Amazon Amazon CloudWatch
- Métriques par défaut pour CloudWatch l'agent avec Amazon EMR
- Configuration de CloudWatch l'agent pour Amazon EMR
- Considérations et restrictions
- CloudWatch historique des versions de l'agent

Créez un cluster EMR qui utilise l'agent Amazon Amazon Amazon CloudWatch

Les procédures de cette section décrivent les étapes de création d'un cluster dans Amazon EMR avec l' CloudWatch agent Amazon à partir de la AWS Management Console et de la. AWS CLI

Rubriques

- Autorisations IAM requises pour l'agent CloudWatch
- Point de terminaison CloudWatch d'agent requis
- Créer un cluster EMR

Autorisations IAM requises pour l'agent CloudWatch

L' CloudWatch agent a besoin de l'`cloudwatch:PutMetricData`autorisation AWS Identity and Access Management (IAM) dans le profil d' EC2 instance Amazon pour Amazon EMR. Le rôle par défaut d’Amazon EMR dispose déjà de cette autorisation. Vous pouvez créer le rôle par défaut à

partir du AWS CLI with`aws emr create-default-roles`. Pour plus d'informations, consultez la section Rôle de service pour les EC2 instances de cluster (profil d'EC2 instance) dans le guide de gestion Amazon EMR.

L'exemple de politique IAM suivant inclut l'autorisation `cloudwatch:PutMetricData`.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "cloudwatch:PutMetricData",
      "Resource": "*"
    }
  ]
}
```

Point de terminaison CloudWatch d'agent requis

Pour publier des métriques CloudWatch pour un cluster EMR dans un sous-réseau privé, créez un point de terminaison d' CloudWatch agent et associez-le au VPC dans lequel se trouve le sous-réseau privé.

Pour plus d'informations sur les CloudWatch points de terminaison de chacun Région AWS, consultez la section CloudWatch Points de terminaison et quotas Amazon dans le Guide de référence AWS général.

Créer un cluster EMR

Une fois que vous avez configuré les autorisations et le point de terminaison requis pour une utilisation avec l' CloudWatch agent, utilisez la AWS Management Console ou la AWS CLI pour créer un cluster avec l'agent installé.

Console

Pour créer un cluster avec l' CloudWatch agent Amazon Amazon depuis la console

1. Accédez à la console Amazon EMR.
2. Choisissez Créer un cluster.
3. Sous Nom et applications, choisissez Amazon EMR 7.0.0 ou version ultérieure.

4. Sous Offre d'applications, sélectionnez l'offre ou les applications à installer sur votre cluster, et incluez CloudWatch l'agent dans votre sélection.
5. Procédez à la création du cluster adapté à votre cas d'utilisation.

AWS CLI

Dans le AWS CLI, vous pouvez ajouter CloudWatch l'agent Amazon à un cluster avec le `--applications` paramètre pour`create-cluster`.

Pour créer un cluster avec l' CloudWatch agent Amazon à partir du AWS CLI

- Lorsque vous créez un cluster, utilisez une commande similaire à ce qui suit pour inclure l' CloudWatch agent Amazon. *myKey*Remplacez-le par le nom de votre paire de EC2 clés.

```
aws emr create-cluster --name "Spark cluster with CloudWatch agent" \
--release-label emr-7.0.0 \
--applications Name=Spark Name=AmazonCloudWatchAgent \
--ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \
--instance-count 3 --use-default-roles
```

Pour plus d'informations sur l'utilisation d'Amazon EMR avec la AWS CLI, consultez la page AWS CLI Command Reference.

Métriques par défaut pour CloudWatch l'agent avec Amazon EMR

Lorsque vous installez l' CloudWatch agent Amazon sur Amazon EMR, la configuration par défaut publie les métriques système suivantes pour toutes les instances de votre cluster, sauf si vous configurez l'agent différemment. Pour les définitions de chaque métrique, consultez la section Mesures collectées par l' CloudWatch agent dans le guide de CloudWatch l'utilisateur Amazon.

CPU

Métriques de processeur

- `cpu_usage_active`
- `cpu_usage_guest`
- `cpu_usage_guest_nice`

- cpu_usage_idle
- cpu_usage_iowait
- cpu_usage_irq
- cpu_usage_nice
- cpu_usage_softirq
- cpu_usage_steal
- cpu_usage_system
- cpu_usage_user

Disk

Métriques de disque

- disk_free
- disk_total
- disk_used
- disk_used_percent

Memory

Métrique de mémoire

- mem_active
- mem_available
- mem_available_percent
- mem_free
- mem_inactive
- mem_total
- mem_used
- mem_used_percent
- mem_buffered
- mem_cached

Network IO

Métriques d’E/S réseau

- `net_bytes_recv`
- `net_bytes_sent`
- `net_packets_recv`
- `net_packets_sent`

Process

Métriques de traitement

- `processes_running`
- `processes_total`

Swap

Métriques d’échange

- `swap_free`
- `swap_used`
- `swap_used_percent`

Par défaut, l'agent publie toutes les métriques CloudWatch dans l'espace de noms `CWAgent` personnalisé et sous le schéma`cluster.id`,, `instance.id``node.type`,`service.name`. Pour consulter ces métriques par défaut publiées par l' CloudWatch agent pour Amazon EMR, procédez comme suit :

1. Accédez à la CloudWatch console Amazon.
2. Sélectionnez l’onglet Métriques, puis Toutes les métriques.
3. Sous Espaces de noms personnalisés, sélectionnez **`CWAgent`**. Sélectionnez ensuite le schéma : **`cluster.id`**,**`instance.id`**,**`node.type`**, **`service.name`**.
4. Continuez à interroger les métriques depuis l' CloudWatch interface en fonction de votre cas d'utilisation.

Amazon EMR attache les étiquettes suivantes à chaque métrique :. `["cluster.id", "instance.id', "node.type", "service.name"]` Tenez compte des éléments suivants en ce qui concerne ces étiquettes :

- La valeur de l'étiquette `cluster.id` correspond à l'ID du cluster EMR qui a produit la métrique. `j-123456789ABC` est un exemple de valeur pour l'étiquette `cluster.id`.
- La valeur de l'étiquette `instance.id` correspond à l'ID de l'instance du cluster EMR qui a produit la métrique. `i-01bcf5f140f3355777` est un exemple de valeur pour l'étiquette `instance.id`.
- La valeur de l'`node.type`étiquette représente le type de nœud de l'instance du cluster qui a produit la métrique. Cette valeur peut être soit `mastercore`, soit`task`.
- La valeur de l'étiquette `service.name` correspond au nom du service qui a produit la métrique. La valeur de `service.name` par défaut pour les métriques système par défaut est `system`.

> **Note**
>
> Amazon EMR R version 7.0.0 utilise un schéma d'étiquette métrique légèrement différent. `cluster.id`est `jobflow.id` et `node.type` n'est pas disponible. `instance.id`et `service.name` sont disponibles.

Configuration de CloudWatch l'agent pour Amazon EMR

Amazon EMR R 7.0.0 et versions ultérieures incluent l'agent Amazon EMR. CloudWatch Les sections suivantes décrivent comment configurer l'agent sur votre cluster EMR comme décrit dans. Créez un cluster EMR qui utilise l'agent Amazon Amazon Amazon CloudWatch

Rubriques

- Configurer CloudWatch l'agent pour Amazon EMR 7.1.0
- Configurer CloudWatch l'agent pour Amazon EMR 7.0.0

Configurer CloudWatch l'agent pour Amazon EMR 7.1.0

À partir d'Amazon EMR 7.1.0, vous pouvez configurer l' CloudWatch agent Amazon pour utiliser des métriques système supplémentaires, ajouter des métriques d'application et modifier la destination des métriques à l'aide de l'API de configuration Amazon EMR. Pour plus d'informations sur l'utilisation

de l'API de configuration EMR pour configurer les applications de votre cluster, consultez la section Configurer les applications.

> **Note**
>
> La version 7.1.0 ne prend en charge que le type de reconfiguration. `OVERWRITE` Pour plus d'informations sur les types de reconfiguration, consultez la section Considérations relatives à la reconfiguration d'un groupe d'instances.

Rubriques

- Schéma de configuration
- Exemples de configurations des métriques système
- Métriques d'application exemples de configurations des métriques
- Exemple d'Amazon Managed Service for Prometheus

Schéma de configuration

`emr-metrics`possède les classifications suivantes :

- `emr-system-metrics`: configurez les métriques système, comme le processeur, le disque et la mémoire.
- `emr-hadoop-hdfs-datanode-metrics`— configurer les métriques Hadoop JMX DataNode
- `emr-hadoop-hdfs-namenode-metrics`— configurer les métriques Hadoop JMX NameNode
- `emr-hadoop-yarn-nodemanager-metrics`— configure les métriques Yarn NodeManager JMX
- `emr-hadoop-yarn-resourcemanager-metrics`— configure les métriques Yarn ResourceManager JMX
- `emr-hbase-master-metrics`— configurer les HBase métriques JMX principales
- `emr-hbase-region-server-metrics`— configurer les métriques JMX de HBase Region Server
- `emr-hbase-rest-server-metrics`— configurer les métriques JMX du serveur HBase REST
- `emr-hbase-thrift-server-metrics`— configurer les métriques HBase JMX de Thrift Server

Les tableaux suivants décrivent les propriétés et les configurations disponibles pour toutes les classifications.

propriétés emr-metrics

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`metrics_destination`	Facultatif	Détermine si les métriques de cluster sont publiées sur Amazon CloudWatch ou Amazon Managed Service for Prometheus.	« SURVEILLANCE DES NUAGES »	« CLOUDWATCH », « PROMÉTHÉE	Cette propriété n'est pas sensible à la casse. Par exemple, « Cloudwatch » est identique à « CLOUDWATCH ».
`prometheus_endpoint`	Facultatif	Si elle `metrics_destination` est définie sur « PROMETHEUS », cette propriété configure l'CloudWatch agent pour qu'il envoie des métriques au point de terminaison d'écriture à distance Amazon Managed	N/A	Toute URL d'écriture à distance Amazon Managed Service for Prometheus valide. Le format de l'URL d'écriture à distance est `https://aps-workspaces.<region>.com/workspaces/<workspac`	Ce champ est obligatoire s'il `metrics_destination` est défini sur « PROMETHEUS ». Le provisionnement échouera si vous ne fournissez pas de clé ou si la valeur est une chaîne vide.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
		Service for Prometheus fourni.		`e_id> / api/v1/r emote_wri te`	

emr-system-metrics propriétés

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`metrics_c ollection _interval`	Facultatif	Fréquence en quelques secondes à laquelle les métriques sont collectée s et publiées par l' CloudWatch agent.	« 60"	Chaîne spécifiant le nombre de secondes. Accepte uniquement les nombres entiers.	Vous pouvez remplacer cette propriété par la `metrics_c ollection _interval` propriété provenant de groupes de mesures individuels.

emr-system-metrics configurations

cpu

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`metrics`	Facultatif	Liste des métriques du	Consultez la section	Liste séparée par	Spécifier une chaîne

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
		processeur que l'agent doit collecter.	Métriques par défaut pour CloudWatch l'agent avec Amazon EMR	des virgules de noms de métriques de processeur valides avec ou sans le `cpu_` préfixe, tel que et. `usage_active` `cpu_time_idle` Consultez la section Mesures collectées par l' CloudWatch agent pour connaître les mesures valides.	vide signifie ne publier aucune métrique du processeur.
`metrics_collection_interval`	Facultatif	Fréquence en secondes à laquelle l'agent doit collecter et publier les métriques du processeur.	La valeur du global`metrics_collection_interval` .	Chaîne spécifiant le nombre de secondes. Accepte uniquement les nombres entiers.	Cette valeur remplace la `metrics_collection_interval` propriété globale uniquement pour les métriques du processeur.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`drop_original_metrics`	Facultatif	Liste des mesures du processeur pour lesquelles il ne faut pas publier de mesures non agrégées.	Aucune métrique non agrégée du processeur n'a été publiée.	Une liste CSV de métriques du processeur qui sont également spécifiées dans la propriété métriques. Une chaîne vide signifie que toutes les métriques du processeur doivent être publiées.	L' CloudWatch agent agrège toutes les métriques par ID de cluster, ID d'instance, type de nœud et nom de service. Par défaut, l' CloudWatch agent ne publie pas les métriques par ressource pour les métriques comportant plusieurs ressources.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`resources`	Facultatif	Détermine si l'agent publiera des métriques par cœur.	"*"	« * » active les métriques par cœur. « » désactive les métriques par cœur.	L' CloudWatch agent publie uniquement les métriques par cœur pour les métriques du processeur qui ne sont pas enregistrées. `drop_original_metrics`

disk

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`metrics`	Facultatif	Liste des métriques de disque que l'agent doit collecter.	Consultez la section Métriques par défaut pour CloudWatch l'agent avec Amazon EMR	Liste séparée par des virgules de noms de métriques de disque valides avec ou sans le `disk_` préfixe, tel que et. `disk_tota`	Si vous spécifiez une chaîne vide, vous ne publiez aucune métrique de disque.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
				`l` `used_perc` `ent` Consultez la section Mesures collectées par l' CloudWatch agent pour connaître les mesures valides.	
`metrics_collection_interval`	Facultatif	Fréquence en secondes à laquelle l'agent doit collecter et publier les métriques du disque.	La valeur du global`metrics` `ollection` `_interval` .	Chaîne spécifiant le nombre de secondes. Accepte uniquement les nombres entiers.	Cette valeur remplace la `metrics_collection_interval` propriété globale uniquement pour les métriques de disque.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`drop_original_metrics`	Facultatif	Liste des mesures de disque pour lesquelles il ne faut pas publier de mesures non agrégées.	Aucune métrique de disque non agrégée n'a été publiée.	Une liste CSV de métriques de disque qui sont également spécifiées dans la propriété métriques. Une chaîne vide signifie que toutes les métriques du disque doivent être publiées.	L' CloudWatch agent agrège toutes les métriques par ID de cluster, ID d'instance, type de nœud et nom de service. Par défaut, l' CloudWatch agent ne publie pas les métriques par ressource pour les métriques comportant plusieurs ressources.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`resources`	Facultatif	Détermine si l'agent publiera per-mount -point des métriques.	"*"	« * » signifie tous les points de montage, « » signifie aucun point de montage ou une liste de points de montage séparés par des virgules. Par exemple, `"/,/emr"`.	L' CloudWatch agent publie uniquement per-mount -point les métriques de disque qui ne sont pas enregistréesdrop_original_metrics .

diskio

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`metrics`	Facultatif	Liste des métriques d'E/S de disque que l'agent doit collecter.	Consultez la section Métriques par défaut pour CloudWatch l'agent avec Amazon EMR	Liste séparée par des virgules de noms de métriques d'E/S de disque valides avec ou sans le `diskio_` préfixe, tel que et. `diskio_re`	Si vous spécifiez une chaîne vide, vous ne publiez aucune métrique d'E/S sur disque.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
				`ads` `writes` Consultez la section Mesures collectées par l' CloudWatch agent pour connaître les mesures valides.	
`metrics_collection_interval`	Facultatif	Fréquence en secondes à laquelle l'agent doit collecter et publier les métriques d'E/S du disque.	La valeur du global`metrics_collection_interval` .	Chaîne spécifiant le nombre de secondes. Accepte uniquement les nombres entiers.	Cette valeur remplace la `metrics_collection_interval` propriété globale uniquement pour les métriques d'E/S de disque.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`drop_original_metrics`	Facultatif	Liste des mesures d'E/S sur disque pour lesquelles il ne faut pas publier de mesures non agrégées.	Aucune métrique d'E/S de disque non agrégée n'a été publiée.	Liste séparée par des virgules de métriques d'E/S sur disque qui sont également spécifiées dans la propriété metrics. Une chaîne vide signifie que toutes les métriques d'E/S du disque doivent être publiées.	L' CloudWatch agent agrège toutes les métriques par ID de cluster, ID d'instance, type de nœud et nom de service. Par défaut, l' CloudWatch agent ne publie pas les métriques par ressource pour les métriques comportant plusieurs ressources.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`resources`	Facultatif	Détermine si l'agent publiera des métriques par appareil.	"*"	« * » désigne tous les périphériques de stockage, « » signifie aucun périphérique de stockage ou une liste de noms de périphériques séparés par des virgules. Par exemple, `"nvme0n1, nvme1n1"` .	L' CloudWatch agent publie uniquement les métriques par appareil pour les métriques d'E/S sur disque qui ne sont pas introduites. `drop_original_metrics`

mem

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`metrics`	Facultatif	Liste des métriques de mémoire que l'agent doit collecter.	Consultez la section Métriques par défaut pour CloudWatch l'agent avec Amazon EMR	Liste séparée par des virgules de noms de métriques de mémoire valides avec ou sans le `mem_` préfixe,	Spécifier une chaîne vide signifie ne publier aucune métrique de mémoire.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
				tels que et. `mem_available` `available_percent` Consultez la section Mesures collectées par l' CloudWatch agent pour connaître les mesures valides.	
`metrics_collection_interval`	Facultatif	Fréquence en secondes à laquelle l'agent doit collecter et publier les métriques de mémoire.	La valeur du global`metrics_collection_interval` .	Chaîne spécifiant le nombre de secondes. Accepte uniquement les nombres entiers.	Cette valeur remplace la `metrics_collection_interval` propriété globale uniquement pour les métriques de mémoire.

net

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`metrics`	Facultatif	Liste des métriques réseau que l'agent doit collecter.	Consultez la section Métriques par défaut pour CloudWatch l'agent avec Amazon EMR	Liste séparée par des virgules de noms de métriques réseau valides avec ou sans le `net_` préfixe, tels que et. `net_packets_sent` `packets_recv` Consultez la section Mesures collectées par l' CloudWatch agent pour connaître les mesures valides.	Si vous spécifiez une chaîne vide, vous ne publiez aucune métrique réseau.
`metrics_collection_interval`	Facultatif	Fréquence en secondes à laquelle l'agent doit collecter et publier les	La valeur du global`metrics_collection_interval` .	Chaîne spécifiant le nombre de secondes. Accepte uniquement	Cette valeur remplace la `metrics_collection_interval` propriété globale

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
		métriques du réseau.		les nombres entiers.	uniquement pour les métriques du réseau.
`drop_original_metrics`	Facultatif	Liste des mesures du réseau pour lesquelles il ne faut pas publier de mesures non agrégées.	Aucune métrique de réseau non agrégée n'a été publiée.	Liste CSV des métriques réseau qui sont également spécifiées dans la propriété métriques. Une chaîne vide signifie que toutes les métriques du réseau doivent être publiées.	L' CloudWatch agent agrège toutes les métriques par ID de cluster, ID d'instance, type de nœud et nom de service. Par défaut, l' CloudWatch agent ne publie pas les métriques par ressource pour les métriques comportant plusieurs ressources.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`resources`	Facultatif	Détermine si l'agent publiera des métriques par interface.	"*"	« * » signifie toutes les interfaces réseau, « » signifie aucune interface réseau ou une liste de noms d'interfaces séparés par des virgules. Par exemple, `"eth0,eth1"` .	L' CloudWatch agent publie uniquement les métriques par interface pour les métriques réseau qui ne sont pas intégrées. `drop_original_metrics`

netstat

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`metrics`	Facultatif	Liste des statistiques réseau que l'agent doit collecter.	Consultez la section Métriques par défaut pour CloudWatch l'agent avec Amazon EMR	Liste séparée par des virgules de noms de métriques de mémoire valides avec ou sans le `netstat_` préfixe, tels que et.	Si vous spécifiez une chaîne vide, vous ne publiez aucune métrique statistique du réseau.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
				`tcp_listen` `netstat_udp_socket` Consultez la section Mesures collectées par l'CloudWatch agent pour connaître les mesures valides.	
`metrics_collection_interval`	Facultatif	Fréquence en secondes à laquelle l'agent doit collecter et publier les statistiques du réseau.	La valeur du global`metrics_collection_interval`.	Chaîne spécifiant le nombre de secondes. Accepte uniquement les nombres entiers.	Cette valeur remplace la `metrics_collection_interval` propriété globale uniquement pour les métriques statistiques du réseau.

processes

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`metrics`	Facultatif	Liste des métriques de processus que l'agent doit collecter.	Consultez la section Métriques par défaut pour CloudWatch l'agent avec Amazon EMR	Liste séparée par des virgules de noms de métriques de mémoire valides avec ou sans le `processes_` préfixe, tels que et. `processes_running` `total` Consultez la section Mesures collectées par l' CloudWatch agent pour connaître les mesures valides.	Spécifier une chaîne vide signifie ne publier aucune métrique de processus.
`metrics_collection_interval`	Facultatif	Fréquence en secondes à laquelle l'agent doit collecter et publier les métriques	La valeur du global`metrics_collection_interval` .	Chaîne spécifiant le nombre de secondes. Accepte uniquement	Cette valeur remplace la `metrics_collection_interval` propriété globale

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
		des processus du système.		les nombres entiers.	uniquement pour les métriques de processus du système.

swap

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
`metrics`	Facultatif	La liste des métriques de swap que l'agent doit collecter.	Consultez la section Métriques par défaut pour CloudWatch l'agent avec Amazon EMR	Liste séparée par des virgules de noms de métriques de mémoire valides avec ou sans le `swap_` préfixe, tels que et. `swap_free` `used_percent` Consultez la section Mesures collectées par l' CloudWatch agent pour	Spécifier une chaîne vide signifie ne publier aucune métrique de swap.

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	Remarques
				connaître les mesures valides.	
`metrics_collection_interval`	Facultatif	Fréquence en secondes à laquelle l'agent doit collecter et publier les métriques de swap.	La valeur du global`metrics_collection_interval` .	Chaîne spécifiant le nombre de secondes. Accepte uniquement les nombres entiers.	Cette valeur remplace la `metrics_collection_interval` propriété globale uniquement pour les métriques de swap.

emr-hadoop-hdfs-datanode-propriétés des métriques

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
`<custom_bean_name>`	Facultatif	N/A	L' CloudWatch agent à partir duquel MBean cet agent doit collecter des métriques , telles queHadoop:service=DataNode,name=DataNod	Chaîne contenant la liste CSV des métriques associées à la. MBean Par exemple, `BlocksCached,BlocksRead` .	

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
			`eActivity` . Vous trouverez des exemples de MBean noms et les métriques correspon dantes dans les exemples de fichiers YAML JMX pour Amazon EMR version 7.0.		
`otel.metric.export.interval`	Facultatif	Fréquence, en milliseco ndes, de collecte des métriques Hadoop. DataNode	« 60000 »	Chaîne spécifiant le nombre de milliseco ndes. Accepte uniquement les nombres entiers.	

emr-hadoop-hdfs-namenode-propriétés des métriques

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
<custom_bean_name>	Facultatif	N/A	L' CloudWatch agent à partir duquel	Chaîne contenant la liste CSV des	

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
			MBean cet agent doit collecter des métriques , telles queHadoop:service=NameNode,name=FSNamesystem . Vous trouverez des exemples de MBean noms et les métriques correspondantes dans les exemples de fichiers YAML JMX pour Amazon EMR version 7.0.	métriques associées à la. MBean Par exemple, BlockCapacity,CapacityUsedGB .	
otel.metric.export.interval	Facultatif	Fréquence, en millisecondes, de collecte des métriques Hadoop. NameNode	« 60000 »	Chaîne spécifiant le nombre de millisecondes. Accepte uniquement les nombres entiers.	

emr-hadoop-yarn-nodemanager-propriétés des métriques

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
<custom_bean_name>	Facultatif	N/A	L' CloudWatch agent à partir duquel MBean cet agent doit collecter des métriques , telles queHadoop:service=NodeManager,name=NodeManagerMetrics . Vous trouverez des exemples de MBean noms et les métriques correspondantes dans les exemples de fichiers YAML JMX pour Amazon EMR version 7.0.	Chaîne contenant la liste CSV des métriques associées à la. MBean Par exemple, `MaxCapacity,AllocatedGB` .	
`otel.metric.export`	Facultatif	Fréquence, en millisecondes, de	« 60000 »	Chaîne spécifiant le nombre	

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
`.interval`		collecte des métriques Hadoop YARN. NodeManager		de millisecondes. Accepte uniquement les nombres entiers.	

emr-hadoop-yarn-resourcemanager-propriétés des métriques

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
<custom_bean_name>	Facultatif	N/A	L' CloudWatch agent à partir duquel MBean cet agent doit collecter des métriques , telles queHadoop:service=ResourceManager,name=PartitionQueueMetrics . Vous trouverez des exemples de MBean noms et les métriques correspon	Chaîne contenant la liste CSV des métriques associées à la. MBean Par exemple, `MaxCapacity,MaxCapacityVCores` .	

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
			dantes dans les exemples de fichiers YAML JMX pour Amazon EMR version 7.0.		
`otel.metric.export.interval`	Facultatif	Fréquence, en millisecondes, de collecte des métriques Hadoop YARN. ResourceManager	« 60000 »	Chaîne spécifiant le nombre de millisecondes. Accepte uniquement les nombres entiers.	

emr-hbase-master-metrics propriétés

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
<custom_bean_name>	Facultatif	N/A	L' CloudWatch agent à partir duquel MBean cet agent doit collecter des métriques , telles queHadoop:service=HBase,name=M	Chaîne contenant la liste CSV des métriques associées à la. MBean Par exemple, `AssignFailedCount, AssignSub`	

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
			`aster,sub=AssignmentManager` . Vous trouverez des exemples de MBean noms et les métriques correspondantes dans les exemples de fichiers YAML JMX pour Amazon EMR version 7.0.	`mittedCount` .	
`otel.metric.export.interval`	Facultatif	Fréquence, en millisecondes, de collecte HBase des métriques principales.	« 60000 »	Chaîne spécifiant le nombre de millisecondes. Accepte uniquement les nombres entiers.	

emr-hbase-region-server-propriétés des métriques

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
<custom_bean_name>	Facultatif	N/A	L' CloudWatch agent à partir duquel MBean cet agent doit collecter des métriques , telles queHadoop:service=HBase,name=RegionServer,sub=IPC . Vous trouverez des exemples de MBean noms et les métriques correspondantes dans les exemples de fichiers YAML JMX pour Amazon EMR version 7.0.	Chaîne contenant la liste CSV des métriques associées à la. MBean Par exemple, numActiveHandler,numActivePriorityHandler .	
otel.metric.export.interval	Facultatif	Fréquence, en millisecondes, de collecte des métriques	« 60000 »	Chaîne spécifiant le nombre de millisecondes.	

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
		du serveur HBase régional.		Accepte uniquement les nombres entiers.	

emr-hbase-rest-server-propriétés des métriques

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
<custom_bean_name>	Facultatif	N/A	L' CloudWatch agent à partir duquel MBean cet agent doit collecter des métriques , telles queHadoop:service=HBase,name=REST . Vous trouverez des exemples de MBean noms et les métriques correspondantes dans les exemples de fichiers YAML JMX pour Amazon	Chaîne contenant la liste CSV des métriques associées à la. MBean Par exemple, `successfulPut,successfulScanCount` .	

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
			EMR version 7.0.		
`otel.metric.export.interval`	Facultatif	Fréquence, en millisecondes, de collecte des métriques du serveur HBase Rest.	« 60000 »	Chaîne spécifiant le nombre de millisecondes. Accepte uniquement les nombres entiers.	

emr-hbase-thrift-server-propriétés des métriques

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
<custom_bean_name>	Facultatif	N/A	L' CloudWatch agent à partir duquel MBean cet agent doit collecter des métriques, telles queHadoop:service=HBase,name=Thrift,sub=ThriftOne . Vous trouverez des	Chaîne contenant la liste CSV des métriques associées à la. MBean Par exemple, `BatchGet_max,BatchGet_mean` .	

Propriété	Obligatoire	Description	Valeur par défaut	Valeurs possibles	
			exemples de MBean noms et les métriques correspon dantes dans les exemples de fichiers YAML JMX pour Amazon EMR version 7.0.		
`otel.metric.export.interval`	Facultatif	Fréquence, en milliseco ndes, de collecte des métriques du serveur HBase Thrift.	« 60000 »	Chaîne spécifiant le nombre de milliseco ndes. Accepte uniquement les nombres entiers.	

Exemples de configurations des métriques système

L'exemple suivant montre comment configurer l' CloudWatch agent pour arrêter d'exporter toutes les métriques système.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-system-metrics",
```

```
        "Properties": {},
        "Configurations": []
      }
    ]
  }
]
```

L'exemple suivant configure l' CloudWatch agent pour exporter les métriques système par défaut. Cela constitue un moyen rapide de réinitialiser l'agent pour qu'il n'exporte les métriques système par défaut que si vous les avez déjà reconfigurées au moins une fois. Cette réinitialisation supprime également toutes les métriques d'application qui ont été reconfigurées auparavant.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": []
  }
]
```

L'exemple suivant configure le cluster pour exporter les `disk` métriques `cpumem`,, et.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-system-metrics",
        "Properties": {
          "metrics_collection_interval": "20"
        },
        "Configurations": [
          {
            "Classification": "cpu",
            "Properties": {
              "metrics": "cpu_usage_guest,cpu_usage_idle",
              "metrics_collection_interval": "30",
              "drop_original_metrics": "cpu_usage_guest"
            }
          },
          {
```

```
            "Classification": "mem",
            "Properties": {
              "metrics": "mem_active"
            }
          },
           {
            "Classification": "disk",
            "Properties": {
              "metrics": "disk_used_percent",
              "resources": "/,/mnt",
              "drop_original_metrics": ""
            }
          }
        ]
      }
    ]
  }
]
```

L'exemple de configuration précédent comporte les propriétés suivantes :

- Toutes les 30 secondes, l'agent collecte la `cpu_guest` métrique pour tous CPUs. Vous pouvez trouver la métrique agrégée sous l'espace de CloudWatch noms`CWAgent > cluster.id, instance.id, node.type, service.name`.
- Toutes les 30 secondes, l'agent collecte la `cpu_idle` métrique pour tous CPUs. Vous pouvez trouver la métrique agrégée sous l'espace de CloudWatch noms`CWAgent > cluster.id, instance.id, node.type, service.name`. L'agent collecte également les métriques par processeur. Vous les trouverez dans le même espace de noms. L'agent collecte cette métrique parce que la `drop_original_metrics` propriété ne contient pas`cpu_idle`, il n'ignore donc pas la métrique.
- Toutes les 20 secondes, l'agent collecte la `mem_active` métrique. Vous pouvez trouver la métrique agrégée sous l'espace de CloudWatch noms`CWAgent > cluster.id, instance.id, node.type, service.name`.
- Toutes les 20 secondes, l'agent collecte les `disk_used_percent` métriques pour les montages `/mnt` sur disque `/` et. Vous pouvez trouver les métriques agrégées sous l'espace de CloudWatch noms`CWAgent > cluster.id, instance.id, node.type, service.name`. L'agent collecte également les métriques par montage. Vous les trouverez dans le même espace de noms. L'agent collecte cette métrique parce que la `drop_original_metrics` propriété ne contient pas`disk_used_percent`, il n'ignore donc pas la métrique.

Métriques d'application exemples de configurations des métriques

L'exemple suivant configure l' CloudWatch agent pour arrêter d'exporter les métriques pour le service Hadoop Namenode.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hadoop-hdfs-namenode-metrics",
        "Properties": {},
        "Configurations": []
      }
    ]
  }
]
```

L'exemple suivant configure un cluster pour exporter les métriques d'applications Hadoop.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hadoop-hdfs-namenode-metrics",
        "Properties": {
          "Hadoop:service=NameNode,name=FSNamesystem": "BlockCapacity,CapacityUsedGB",
          "otel.metric.export.interval": "20000"
        },
        "Configurations": []
      },
       {
        "Classification": "emr-hadoop-hdfs-datanode-metrics",
        "Properties": {
          "Hadoop:service=DataNode,name=JvmMetrics": "MemNonHeapUsedM",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      },
       {
```

```
        "Classification": "emr-hadoop-yarn-resourcemanager-metrics",
        "Properties": {
          "Hadoop:service=ResourceManager,name=CapacitySchedulerMetrics":
 "AllocateNumOps,NodeUpdateNumOps"
        },
        "Configurations": []
      }
    ]
  }
]
```

L'exemple précédent comporte les propriétés suivantes :

- Toutes les 20 secondes, l'agent collecte les `CapacityUsedGB` métriques `BlockCapacity` et auprès des instances exécutant le service Hadoop Namenode.
- Toutes les 30 secondes, l'agent collecte les `MemNonHeapUsedM` métriques des instances exécutant le service Hadoop Datanode.
- Toutes les 30 secondes, l'agent collecte les `NodeUpdateNumOps` métriques `AllocateNumOps` et à partir des instances qui exécutent le Hadoop YARN. ResourceManaager

Exemple d'Amazon Managed Service for Prometheus

L'exemple suivant montre comment configurer l' CloudWatch agent pour exporter des métriques vers Amazon Managed Service for Prometheus.

Si vous exportez actuellement des métriques vers Amazon Managed Service for Prometheus et que vous souhaitez reconfigurer les métriques pour le cluster et continuer à exporter des métriques vers Amazon Managed Service for Prometheus, vous devez inclure les propriétés et. `metrics_destination` `prometheus_endpoint`

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "prometheus",
      "prometheus_endpoint": "http://amp-workspace/api/v1/remote_write"
    },
    "Configurations": []
  }
]
```

Pour utiliser l' CloudWatch agent vers lequel exporter des métriques CloudWatch, utilisez l'exemple suivant.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "cloudwatch"
    },
    "Configurations": []
  }
]
```

> **Note**
>
> L' CloudWatch agent dispose d'un exportateur Prometheus qui renomme certains attributs. Pour les étiquettes de métrique par défaut, Amazon Managed Service for Prometheus utilise des caractères de soulignement à la place des points utilisés par Amazon. CloudWatch Si vous utilisez Amazon Managed Grafana pour visualiser les métriques par défaut dans Amazon Managed Service for Prometheus, les étiquettes s'affichent sous la forme,, et. `cluster_id` `instance_id` `node_type` `service_name`

Configurer CloudWatch l'agent pour Amazon EMR 7.0.0

Vous pouvez configurer l' CloudWatch agent Amazon pour utiliser des métriques système supplémentaires en plus de celles fournies par la configuration par défaut de l' CloudWatch agent. La configuration de la version 7.0.0 requiert l'utilisation d'actions d'amorçage, dont nous avons fourni des exemples dans les sections suivantes. Dans une prochaine version, Amazon EMR fournira des options de configuration supplémentaires via l'API Amazon EMR.

Rubriques

- Configurer des métriques système supplémentaires avec Amazon EMR 7.0.0
- Configurer les métriques d'application avec Amazon EMR 7.0.0
- Configurer Amazon Managed Service for Prometheus en tant que stockage cloud pour les métriques avec Amazon EMR 7.0.0

Configurer des métriques système supplémentaires avec Amazon EMR 7.0.0

Pour configurer l'agent afin d'utiliser un ensemble différent de métriques système dans Amazon EMR 7.0.0, procédez comme suit :

1. Créez ou choisissez un compartiment dans votre compte Amazon S3 dans lequel vous souhaitez stocker les fichiers de configuration qui spécifient les métriques de l' CloudWatch agent.
2. Créez le fichier de configuration `emr-amazon-cloudwatch-agent.json` en spécifiant vos métriques préférées. Pour ce faire, procédez selon l'une des méthodes décrites dans la rubrique Créer le fichier de configuration de l' CloudWatch agent. Pour plus d'informations sur la structure du fichier de configuration d' CloudWatch agent, consultez la rubrique Création ou édition manuelle du fichier de configuration d' CloudWatch agent du Guide de CloudWatch l'utilisateur Amazon.
3. Accédez ensuite au aws-emr-utilities référentiel GitHub et téléchargez les scripts de métriques système suivants :

   - **`install_system_metrics_launcher.sh`** – Script qui télécharge et exécute `install_system_metrics.sh` en arrière-plan pour que le nœud puisse terminer l'amorçage.
   - **`install_system_metrics.sh`** – Script qui attend que l'instance qu'il exécute termine l'amorçage, puis télécharge et applique la configuration dans le fichier JSON.
4. Ouvrez chaque fichier SH et remplacez *`amzn-s3-demo-bucket`* par le nom de votre compartiment indiqué à l'étape 1.
5. Chargez un fichier JSON et deux fichiers SH dans votre compartiment S3.
6. Vous pouvez désormais accéder à la console Amazon EMR et créer un cluster avec l' CloudWatch agent. Sous EMR activé EC2 dans le menu de navigation de gauche, sélectionnez Clusters puis Create cluster.
7. Dans la section Nom et applications, choisissez Amazon EMR 7.0.0 ou version ultérieure.
8. Sous Offre d'applications, sélectionnez l'offre ou les applications à installer sur votre cluster, et incluez CloudWatch l'agent Amazon dans votre sélection.
9. Dans la section Actions d'amorçage, sélectionnez Ajouter.

   - Pour le Nom, insérez **`install_system_metrics_launcher.sh`**.
   - Pour l'Emplacement du script, insérez **`s3://`*`amzn-s3-demo-bucket`*`/install_system_metrics_launcher.sh`**. Remplacez *`amzn-s3-demo-bucket`* par le chemin de votre compartiment S3.

- Laissez le bloc Arguments vide.

10. Sélectionnez Ajouter une action d'amorçage.
11. Continuez à créer le cluster pour répondre à vos besoins en matière de charge de travail.

Au lancement de votre cluster, l' CloudWatch agent publie sur CloudWatch.

Configurer les métriques d’application avec Amazon EMR 7.0.0

Vous pouvez configurer l' CloudWatch agent Amazon pour publier des métriques d'application pour HDFS et YARN en plus des métriques système. Pour configurer l’agent de sorte à publier des métriques d’application, procédez comme suit :

1. Créez ou choisissez un compartiment dans votre compte Amazon S3 dans lequel vous souhaitez stocker les fichiers de configuration qui spécifient les métriques de l' CloudWatch agent.
2. Accédez ensuite au aws-emr-utilities référentiel GitHub et téléchargez les scripts suivants :

   - **`install_app_metrics_launcher.sh`** – Script qui télécharge et exécute `install_app_metrics.sh` en arrière-plan pour que le nœud puisse terminer l’amorçage.
   - **`install_app_metrics.sh`** – Script qui attend que l’instance qu’il exécute termine l’amorçage, puis télécharge et applique la configuration dans les fichiers YAML que vous allez télécharger lors d’une prochaine étape.

3. Ouvrez chaque fichier et remplacez *`amzn-s3-demo-bucket`* par le nom de votre compartiment indiqué à l’étape 1.
4. Téléchargez ensuite les fichiers de mappage YAML suivants. Pour plus d'informations sur la structure de ces fichiers YAML, consultez `javaagent`le OpenTelemetry Instrumentation for Java GitHub dépôt.

   - **`datanode-metrics.yaml`** –Configuration des métriques Hadoop DataNode.
   - **`namenode-metrics.yaml`** –Configuration des métriques Hadoop NameNode.
   - **`nodemanager-metrics.yaml`** –Configuration des métriques Yarn NodeManager.
   - **`resourcemanager-metrics.yaml`** –Configuration des métriques Yarn ResourceManager.

5. Chargez les deux fichiers SH et quatre fichiers YAML dans votre compartiment S3.
6. Vous pouvez désormais accéder à la console Amazon EMR et créer un cluster avec l' CloudWatch agent. Sous EMR activé EC2 dans le menu de navigation de gauche, sélectionnez Clusters puis Create cluster.

7. Dans la section Nom et applications, choisissez Amazon EMR 7.0.0 ou version ultérieure.
8. Sous Offre d'applications, sélectionnez l'offre ou le groupe personnalisé d'applications à installer sur votre cluster, et incluez CloudWatch l'agent dans votre sélection.
9. Dans la section Actions d'amorçage, sélectionnez Ajouter.
   - Pour le Nom, insérez **`install_app_metrics_launcher.sh`**.
   - Pour l'Emplacement du script, insérez **`s3://amzn-s3-demo-bucket/install_app_metrics_launcher.sh`**. Remplacez *`amzn-s3-demo-bucket`* par le chemin de votre compartiment S3.
   - Laissez le bloc Arguments vide.
10. Sélectionnez Ajouter une action d'amorçage.
11. Continuez à créer le cluster pour répondre à vos besoins en matière de charge de travail.

Au lancement de votre cluster, l' CloudWatch agent publie les métriques d'application que vous avez spécifiées avec les métriques système CloudWatch.

Configurer Amazon Managed Service for Prometheus en tant que stockage cloud pour les métriques avec Amazon EMR 7.0.0

Vous pouvez configurer l' CloudWatch agent Amazon pour publier des métriques sur Amazon Managed Service for Prometheus au lieu de. CloudWatch

> **Note**
>
> Vous pouvez publier des métriques depuis l' CloudWatch agent Amazon Managed Service for Prometheus ou sur CloudWatch Amazon, mais vous ne pouvez pas les publier sur les deux services pour le même cluster.

Pour configurer l'agent de sorte à publier des métriques sur Amazon Managed Service for Prometheus, vous devez ajouter `aps:RemoteWrite` AWS Identity and Access Management l'autorisation (IAM) au profil d' EC2instance Amazon EMR. L'exemple de politique suivant contient les autorisations requises :

```
{
  "Version": "2012-10-17",
  "Statement": [
```

```
    {
      "Effect": "Allow",
      "Action": "aps:RemoteWrite",
      "Resource": "*"
    }
  ]
}
```

Utiliser l' CloudWatch agent sur un cluster EMR pour publier des métriques sur Amazon Managed Service for Prometheus

Une fois que la politique de service dispose des autorisations appropriées, procédez comme suit pour lancer un cluster qui utilise l' CloudWatch agent pour publier des métriques sur Amazon Managed Service for Prometheus.

1. Utilisez la AWS Management Console ou la AWS CLI pour créer un Workspace Amazon Managed Service for Prometheus. Pour plus d'informations, voir la rubrique Création d'un Workspace du Guide de l'utilisateur Amazon Managed Service for Prometheus.
2. Créez ou choisissez un compartiment dans votre compte Amazon S3 dans lequel vous souhaitez stocker les fichiers de lancement qui spécifient Amazon Managed Service for Prometheus en tant qu'espace de stockage cloud.
3. Accédez ensuite au aws-emr-utilities référentiel GitHub et téléchargez les scripts suivants :

   - **`add_prometheus_endpoint_launcher.sh`** – Script qui télécharge et exécute `add_prometheus_endpoint.sh` en arrière-plan pour que le nœud puisse terminer l'amorçage.
   - **`add_prometheus_endpoint.sh`**— Script qui attend que l'instance qu'il exécute termine l'amorçage, puis configure l' CloudWatch agent à publier sur le point de terminaison Amazon Managed Service for Prometheus que vous spécifiez comme argument lorsque vous lancez votre cluster.
4. Ouvrez chaque fichier et remplacez *`amzn-s3-demo-bucket`* par le nom de votre compartiment indiqué à l'étape 2.
5. Utilisez la AWS CLI pour créer un cluster EMR avec l'action d'`add_prometheus_endpoint_launcher.sh`amorçage. Dans la commande suivante, remplacez *`amzn-s3-demo-bucket`* par le compartiment contenant l'action d'amorçage et remplacez *`managedpro-remote-write-workspace-url`* par le point de terminaison d'écriture à distance pour votre Workspace Amazon Managed Service for Prometheus. Assurez-vous de spécifier une étiquette de version Amazon EMR `emr-7.0.0` ou ultérieure.

```
aws emr create-cluster --name managedpro-cluster \
    --release-label emr-7.0.0 \
    --applications Name=Hadoop Name=AmazonCloudWatchAgent \
    --ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \
    --instance-count 3 --use-default-roles
   --bootstrap-actions Name='Add Prometheus Endpoint',Path=s3://amzn-s3-demo-
bucket/add_prometheus_endpoint_launcher.sh,Args='managedpro-remote-write-workspace-
url'
```

Au lancement de votre cluster, l' CloudWatch agent publie les métriques qu'il collecte sur Amazon Managed Service for Prometheus.

Utiliser Amazon Managed Service for Prometheus en tant que source de données pour Amazon Managed Grafana

Une fois qu'Amazon EMR a publié les métriques du cluster sur Amazon Managed Service for Prometheus, procédez comme suit pour visualiser les métriques avec Amazon Managed Grafana :

1. Utilisez la AWS Management Console pour créer un Workspace et un utilisateur Amazon Managed Grafana disposant des autorisations appropriées. Pour plus d'informations, voir la rubrique Création d'un Workspace du Guide de l'utilisateur Amazon Managed Grafana.
2. Ajoutez votre Workspace Amazon Managed Service for Prometheus en tant que source de données pour Amazon Managed Grafana. Pour plus d'informations, voir la rubrique Utiliser la configuration de la source de données AWS pour ajouter Amazon Managed Service for Prometheus en tant que source de données du Guide de l'utilisateur Amazon Managed Grafana.

> **Note**
>
> L' CloudWatch agent dispose d'un exportateur Prometheus qui renomme certains attributs. Pour les étiquettes de métrique par défaut, Amazon Managed Service for Prometheus utilise des caractères de soulignement à la place des points utilisés par Amazon. CloudWatch Si vous utilisez Amazon Managed Grafana pour visualiser les métriques par défaut dans Amazon Managed Service for Prometheus, les étiquettes s'affichent sous la forme `jobflow_id`, `instance_id` et `service_name`.

De plus, toutes les métriques d'application publiées par l' CloudWatch agent sur Amazon Managed Service for Prometheus utilisent l'étiquette à la place de `job`. `service_name` Cependant, les métriques système continuent d'utiliser l'étiquette `service_name`.

Considérations et restrictions

- L' CloudWatch agent Amazon natif est disponible pour les clusters que vous créez avec Amazon EMR 7.0.0 et versions ultérieures.
- La configuration de l' CloudWatch agent dans Amazon EMR 7.0.0 requiert l'utilisation d'actions d'amorçage. Dans une prochaine version, Amazon EMR fournira des options de configuration supplémentaires via l'API Amazon EMR.
- Vous ne pouvez pas installer l' CloudWatch agent Amazon EMR si vous l'avez déjà déployé à l'aide CloudWatch d'une autre méthode, telle que AWS Systems Manager Agent (SSM Agent) dans la région où vous créez votre cluster. Cette opération entraînera l'arrêt de votre cluster avec des erreurs.
- L' CloudWatch GetMetricData API prend en charge jusqu'à 500 métriques par demande. Si votre cluster Amazon EMR contient plus de 250 nœuds dans un groupe d'instances ou un parc, les graphiques correspondants dans le tableau de bord CloudWatch intégré à EMR apparaîtront vides avec l'erreur Trop de métriques car ces métriques nécessitent deux points de données par métrique dans le tableau de bord d'aperçu du cluster. Toutefois, en filtrant les tableaux de bord des groupes d'instances principaux ou de tâches (ou du parc), vous pourrez afficher les graphiques correspondant à un maximum de 500 nœuds par groupe d'instances ou parc. Cela est dû au fait qu'ils ne nécessitent pas deux points de données par métrique. Au-delà de 500 nœuds par groupe d'instances ou parc, l'erreur Too many metrics se produit également pour les métriques de ces tableaux de bord.

CloudWatch historique des versions de l'agent

Le tableau suivant répertorie les versions AmazonCloudWatchAgent incluses dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

AmazonCloudWatchAgent informations sur la version

Étiquette de version Amazon EMR	AmazonCloudWatchAgent Version	Composants installés avec AmazonCloudWatchAgent
emr-7.9.0	1,300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server
emr-7.8.0	1,300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server
emr-7,7.0	1,300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd

Étiquette de version Amazon EMR	AmazonCloudWatchAgent Version	Composants installés avec AmazonCloudWatchAgent
		fs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.6.0	1,300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	AmazonCloudWatchAgent Version	Composants installés avec AmazonCloudWatchAgent
emr-7.5.0	1,300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server
emr-7.4.0	1,300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	AmazonCloudWatchAgent Version	Composants installés avec AmazonCloudWatchAgent
emr-7.3.0	1,300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server
emr-7.2.0	1,300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	AmazonCloudWatchAgent Version	Composants installés avec AmazonCloudWatchAgent
emr-7.1.0	1,300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.0.0	1,300031.1	adot-java-agent, emr-amazon-cloudwatch-agent

Delta Lake

Delta Lake est un framework de couche de stockage pour les architectures Lakehouse généralement construites sur Amazon S3. Avec les versions 6.9.0 et supérieures d'Amazon EMR, vous pouvez utiliser Apache Spark 3.x sur des clusters Amazon EMR dotés de tables Delta Lake. Pour plus d'informations sur les cabanes au bord du lac Delta, consultez le site https://delta.io/.

Le tableau suivant répertorie la version de Delta incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants qu'Amazon EMR installe avec Delta.

Pour la version des composants installés avec Delta dans cette version, voir la version 7.9.0 Versions des composants.

Informations sur la version Delta pour emr-7.9.0

Étiquette de version Amazon EMR	Version Delta	Composants installés avec Delta
emr-7.9.0	Delta 3.3.0-amzn-1	Not available.

Le tableau suivant répertorie la version de Delta incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Delta.

Pour connaître la version des composants installés avec Delta dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version Delta pour emr-6.15.0

Étiquette de version Amazon EMR	Version Delta	Composants installés avec Delta
emr-6.15.0	Delta 2.4.0	Not available.

Présentation de Delta Lake

Delta Lake est un projet open source qui permet de mettre en œuvre des architectures de lacs de données modernes généralement basées sur Amazon S3. Delta Lake propose les fonctions suivantes :

- Transactions atomiques, cohérentes, isolées et durables (ACID) sur Spark. Les lecteurs bénéficient d'une vue cohérente du tableau lors d'une tâche Spark.
- Gestion évolutive des métadonnées avec traitement distribué par Spark.
- Combine les cas de streaming et d'utilisation par lots avec la même table Delta.
- Application automatique du schéma pour éviter les mauvais enregistrements lors de l'ingestion de données.
- Voyage dans le temps avec gestion des versions des données.
- Prend en charge les opérations de fusion, de mise à jour et de suppression pour les cas d'utilisation complexes tels que la capture des données modifiées (CDC), le streaming upserts, etc.

Utilisation d'un cluster sur lequel Delta Lake est installé

Rubriques

- Utiliser un cluster Delta Lake avec Flink
- Utilisez un cluster Delta Lake avec Trino
- Utiliser un cluster Delta Lake avec Spark
- Utilisez un cluster Delta Lake avec Spark et AWS Glue

Utiliser un cluster Delta Lake avec Flink

Avec Amazon EMR version 6.11 ou ultérieure, vous pouvez utiliser Delta Lake avec votre cluster Flink. Les exemples suivants utilisent le AWS CLI pour travailler avec Delta Lake sur un cluster Amazon EMR Flink.

> **Note**
>
> Amazon EMR prend en charge l' DataStream API Flink lorsque vous utilisez Delta Lake avec un cluster Flink.

Création d'un cluster Delta Lake

1. Créez un fichier, `delta_configurations.json`, contenant les éléments suivants :

```
[{"Classification":"delta-defaults",
```

```
    "Properties":{"delta.enabled":"true"}}]
```

2. Créez un cluster avec la configuration suivante. Remplacez `example Amazon S3 bucket path` et `subnet ID` par les vôtres.

```
aws emr create-cluster
--release-label emr-6.11.0
--applications Name=Flink
--configurations file://delta_configurations.json
--region us-east-1  --name My_Spark_Delta_Cluster
--log-uri  s3://amzn-s3-demo-bucket/
--instance-type m5.xlarge
--instance-count 3
--service-role EMR_DefaultRole_V2
--ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Initialisation d'une session Flink yarn

Pour initialiser une session Flink yarn, exécutez la commande suivante :

```
flink-yarn-session -d
```

Créer une tâche Flink avec Delta Lake

Les exemples suivants montrent comment utiliser sbt ou Maven pour créer votre tâche Flink avec Delta Lake.

sbt

sbt est un outil de construction pour Scala que vous pouvez utiliser avec peu ou pas de configuration lorsque vous avez de petits projets.

```
libraryDependencies ++= Seq(
  "io.delta" %% "delta-flink" % deltaConnectorsVersion % "provided",
  "io.delta" %% "delta-standalone" % deltaConnectorsVersion % "provided",
  "org.apache.flink" %% "flink-clients" % flinkVersion % "provided",
  "org.apache.flink" %% "flink-parquet" % flinkVersion % "provided",
  "org.apache.hadoop" % "hadoop-client" % hadoopVersion % "provided",
  "org.apache.flink" % "flink-table-common" % flinkVersion % "provided",
  "org.apache.flink" %% "flink-table-runtime" % flinkVersion % "provided")
```

Maven

Maven est un outil d'automatisation de build open source de l'Apache Software Foundation. Avec Maven, vous pouvez créer, publier et déployer une tâche Flink avec Delta Lake sur Amazon EMR.

```
<project>
<properties>
    <scala.main.version>2.12</scala.main.version>
    <delta-connectors-version>0.6.0</delta-connectors-version>
    <flink-version>1.16.1</flink-version>
    <hadoop-version>3.1.0</hadoop-version>
</properties>

<dependencies>
    <dependency>
        <groupId>io.delta</groupId>
        <artifactId>delta-flink</artifactId>
        <version>$delta-connectors-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>io.delta</groupId>
        <artifactId>delta-standalone_$scala-main-version</artifactId>
        <version>$delta-connectors-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-clients</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-parquet</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.hadoop</groupId>
        <artifactId>hadoop-client</artifactId>
        <version>$hadoop-version</version>
        <scope>provided</scope>
    </dependency>
```

```
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-table-common</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-table-runtime</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
</dependencies>
```

Écrire dans une table Delta avec l'API Flink Datastream

Utilisez l'exemple suivant pour créer un DeltaSink pour écrire dans la table avec un `deltaTablePath:`

```
public static DataStream<RowData> createDeltaSink(
        DataStream<RowData> stream,
        String deltaTablePath,
        RowType rowType) {
    Configuration configuration = new Configuration();
    DeltaSink<RowData> deltaSink = DeltaSink
            .forRowData(
                    new org.apache.flink.core.fs.Path(deltaTablePath),
                    configuration,
                    rowType)
            .build();
    stream.sinkTo(deltaSink);
    return stream;
}
```

Lire à partir d'une table Delta avec l'API Flink Datastream

Utilisez l'exemple suivant pour créer un borné DeltaSource à lire dans le tableau avec un `deltaTablePath:`

```
public static DataStream<RowData> createBoundedDeltaSourceAllColumns(
        StreamExecutionEnvironment env,
```

```
        String deltaTablePath) {
    Configuration configuration = new Configuration();
    DeltaSource<RowData> deltaSource = DeltaSource
            .forBoundedRowData(
                    new org.apache.flink.core.fs.Path(deltaTablePath),
                    configuration)
            .build();

    return env.fromSource(deltaSource, WatermarkStrategy.noWatermarks(), "delta-
source");
}
```

Création de récepteurs avec prise en charge de plusieurs clusters pour le système autonome de Delta Lake

Utilisez l'exemple suivant pour créer une table DeltaSink à écrire avec un `deltaTablePath` support multi-clusters :

```
public DataStream<RowData> createDeltaSink(
        DataStream<RowData> stream,
        String deltaTablePath) {
    Configuration configuration = new Configuration();
    configuration.set("spark.delta.logStore.s3.impl",
 "io.delta.storage.S3DynamoDBLogStore");
    configuration.set("spark.io.delta.storage.S3DynamoDBLogStore.ddb.tableName",
 "delta_log");
    configuration.set("spark.io.delta.storage.S3DynamoDBLogStore.ddb.region", "us-
east-1");

    DeltaSink<RowData> deltaSink = DeltaSink
        .forRowData(
            new Path(deltaTablePath),
            configuration,
            rowType)
        .build();
    stream.sinkTo(deltaSink);
    return stream;
}
```

Exécuter la tâche Flink

Utilisez la commande suivante pour exécuter votre tâche :

```
flink run FlinkJob.jar
```

Utilisez un cluster Delta Lake avec Trino

Avec les versions 6.9.0 et supérieures d'Amazon EMR, vous pouvez utiliser Delta Lake avec votre cluster Trino.

Dans ce didacticiel, nous allons utiliser le AWS CLI cluster Trino pour travailler avec Delta Lake sur Amazon EMR.

Création d'un cluster Delta Lake

1. Créez un fichier, `delta_configurations.json`, et définissez des valeurs pour le catalogue de votre choix. Par exemple, si vous souhaitez utiliser le métastore Hive comme catalogue, le contenu de votre fichier doit être le suivant :

   ```
   [{"Classification":"delta-defaults",
       "Properties":{"delta.enabled":"true"}},
       {"Classification":"trino-connector-delta",
       "Properties":{"hive.metastore.uri":"thrift://localhost:9083"}}]
   ```

   Si vous souhaitez utiliser le AWS Glue Catalog comme boutique, le contenu de votre fichier doit être le suivant :

   ```
   [{"Classification":"delta-defaults",
       "Properties":{"delta.enabled":"true"}},
       {"Classification":"trino-connector-delta",
       "Properties":{"hive.metastore":"glue"}}]
   ```

2. Créez un cluster avec la configuration suivante, en remplaçant **`example Amazon S3 bucket path`** et **`subnet ID`** par les vôtres.

   ```
   aws emr create-cluster
       --release-label emr-6.9.0
       --applications Name=Trino
       --configurations file://delta_configurations.json
       --region us-east-1  --name My_Spark_Delta_Cluster
       --log-uri  s3://amzn-s3-demo-bucket/
       --instance-type m5.xlarge
       --instance-count 2
   ```

```
    --service-role EMR_DefaultRole_V2
    --ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Initialisation de la session Trino pour Delta Lake

Pour initialiser la session Trino, exécutez la commande suivante

```
trino-cli --catalog delta
```

Écriture dans une table Delta Lake

Créez et écrivez dans votre table à l'aide des commandes SQL suivantes :

```
SHOW SCHEMAS;

CREATE TABLE default.delta_table (id  int, data varchar, category varchar) WITH
( location =  's3://amzn-s3-demo-bucket/<prefix>');

INSERT INTO default.delta_table VALUES  (1,'a','c1'), (2,'b','c2'), (3,'c','c3');
```

Lecture à partir d'une table Delta Lake

Lisez le contenu de votre table à l'aide de la commande SQL suivante :

```
SELECT * from default.delta_table;
```

Utiliser un cluster Delta Lake avec Spark

À partir de la version 6.9.0 d'Amazon EMR, vous pouvez utiliser Delta Lake avec votre cluster Spark sans avoir besoin d'actions d'amorçage. Pour les versions 6.8.0 et antérieures d'Amazon EMR, vous pouvez utiliser les actions amorçage pour préinstaller les dépendances nécessaires.

Les exemples suivants utilisent le AWS CLI pour travailler avec Delta Lake sur un cluster Amazon EMR Spark.

Pour utiliser Delta Lake sur Amazon EMR avec le AWS Command Line Interface, créez d'abord un cluster. Pour plus d'informations sur la façon de spécifier la classification Delta Lake avec AWS Command Line Interface, voir Fournir une configuration à l'aide du AWS Command Line Interface lorsque vous créez un cluster ou Fournir une configuration avec le SDK Java lorsque vous créez un cluster.

1. Créez un fichier, `configurations.json`, contenant les éléments suivants :

```
[{"Classification":"delta-defaults",  "Properties":{"delta.enabled":"true"} }]
```

2. Créez un cluster avec la configuration suivante, en remplaçant l'exemple d'Amazon S3 **`bucket path`** et le **`subnet ID`** par le vôtre.

```
aws emr create-cluster
     --release-label  emr-6.9.0
     --applications Name=Spark
     --configurations file://delta_configurations.json
     --region us-east-1
     --name My_Spark_Delta_Cluster
     --log-uri  s3://amzn-s3-demo-bucket/
     --instance-type m5.xlarge
     --instance-count 2
     --service-role EMR_DefaultRole_V2
     --ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Vous pouvez également créer un cluster Amazon EMR et une application Spark avec les fichiers suivants en tant que dépendances JAR dans une tâche Spark :

```
/usr/share/aws/delta/lib/delta-core.jar,
/usr/share/aws/delta/lib/delta-storage.jar,
/usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar
```

> **Note**
>
> Si vous utilisez les versions 7.0.0 ou supérieures d'Amazon EMR, utilisez `/usr/share/aws/delta/lib/delta-spark.jar` à la place de. `/usr/share/aws/delta/lib/delta-core.jar`

Pour plus d'informations, consultez Soumission d'applications.

Pour inclure une dépendance JAR dans la tâche Spark, vous pouvez ajouter les propriétés de configuration suivantes à l'application Spark :

```
--conf “spark.jars=/usr/share/aws/delta/lib/delta-core.jar,
     /usr/share/aws/delta/lib/delta-storage.jar,
     /usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar"
```

Pour plus d'informations sur les dépendances entre les tâches de Spark, consultez Gestion des dépendances.

Si vous utilisez les versions 7.0.0 ou supérieures d'Amazon EMR, ajoutez plutôt la `/usr/share/aws/delta/lib/delta-spark.jar` configuration.

```
--conf “spark.jars=/usr/share/aws/delta/lib/delta-spark.jar,
     /usr/share/aws/delta/lib/delta-storage.jar,
     /usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar"
```

Initialisation d'une session Spark pour Delta Lake

Les exemples suivants montrent comment lancer le shell interactif Spark, utiliser Spark submit ou utiliser Blocs-notes EMR pour travailler avec Delta Lake sur Amazon EMR.

spark-shell

1. Connexion au nœud primaire à l'aide de SSH. Pour plus d'informations, consultez Connexion au nœud primaire à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
2. Entrez la commande suivante pour lancer le shell Spark. Pour utiliser la PySpark coque, remplacez-la `spark-shell` par`pyspark`.

   ```
   spark-shell \
      --conf "spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension" \
      --conf
    "spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog"
   ```

Si vous exécutez Amazon EMR version 6.15.0 ou ultérieure, vous devez également utiliser les configurations suivantes pour utiliser un contrôle d'accès précis basé sur Lake Formation with Delta Lake.

```
spark-shell \
  --conf
 spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserv
 \
  --conf
 spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog \

  --conf spark.sql.catalog.spark_catalog.lf.managed=true
```

spark-submit

1. Connexion au nœud primaire à l'aide de SSH. Pour plus d'informations, consultez Connexion au nœud primaire à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
2. Entrez la commande suivante pour lancer la session Spark pour Delta Lake.

```
spark-submit
–conf "spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension"
–conf
 "spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog"
```

Si vous exécutez Amazon EMR version 6.15.0 ou ultérieure, vous devez également utiliser les configurations suivantes pour utiliser un contrôle d'accès précis basé sur Lake Formation with Delta Lake.

```
spark-submit \  `
--conf
 spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserv
--conf
 spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog \

--conf spark.sql.catalog.spark_catalog.lf.managed=true
```

EMR Studio notebooks

Pour initialiser une session Spark à l'aide des blocs-notes Amazon EMR Studio, configurez votre session Spark à l'aide de la commande magique %%configure de votre bloc-notes Amazon EMR, comme dans l'exemple suivant. Pour plus d'informations, consultez Utilisation des magies de Blocs-notes EMR dans le Guide de gestion Amazon EMR.

```
%%configure -f
{
  "conf": {
    "spark.sql.extensions":  "io.delta.sql.DeltaSparkSessionExtension",
     "spark.sql.catalog.spark_catalog":
 "org.apache.spark.sql.delta.catalog.DeltaCatalog"
  }
}
```

Si vous exécutez Amazon EMR version 6.15.0 ou ultérieure, vous devez également utiliser les configurations suivantes pour utiliser un contrôle d'accès précis basé sur Lake Formation with Delta Lake.

```
%%configure -f
{
  "conf": {
    "spark.sql.extensions":
 "io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserver.connector.spark.sql
    "spark.sql.catalog.spark_catalog":
 "org.apache.spark.sql.delta.catalog.DeltaCatalog",
    "spark.sql.catalog.spark_catalog.lf.managed": "true"
  }
}
```

Écriture dans une table Delta Lake

L'exemple suivant montre comment créer DataFrame et écrire un jeu de données Delta Lake. L'exemple montre comment travailler avec des jeux de données avec le shell Spark lorsque vous êtes connecté au nœud primaire en utilisant SSH comme utilisateur hadoop par défaut.

> **Note**
>
> Pour coller des exemples de code dans le shell Spark, tapez :paste à l'invite, collez l'exemple, puis appuyez sur CTRL + D.

PySpark

Spark inclut un shell basé sur Python, `pyspark`, que vous pouvez utiliser pour créer un prototype de programmes Spark écrits en Python. Tout comme pour `spark-shell`, invoquez `pyspark` sur le nœud primaire.

```
## Create a DataFrame
data =  spark.createDataFrame([("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
("101",  "2015-01-01", "2015-01-01T12:14:58.597216Z"),
("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103",  "2015-01-01",  "2015-01-01T13:51:40.519832Z")],
["id", "creation_date",  "last_update_time"])

## Write a DataFrame as a Delta Lake dataset to the S3  location
spark.sql("""CREATE  TABLE IF NOT EXISTS delta_table (id string, creation_date
 string,
last_update_time string)
USING delta location
's3://amzn-s3-demo-bucket/example-prefix/db/delta_table'""");

data.writeTo("delta_table").append()
```

Scala

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._

// Create a DataFrame
val data = Seq(("100",  "2015-01-01",  "2015-01-01T13:51:39.340396Z"),
("101",  "2015-01-01",  "2015-01-01T12:14:58.597216Z"),
("102",  "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103",  "2015-01-01",  "2015-01-01T13:51:40.519832Z")).toDF("id", "creation_date",
 "last_update_time")

// Write a DataFrame as a Delta Lake dataset to the S3  location
```

```
spark.sql("""CREATE  TABLE IF NOT EXISTS delta_table (id string,
creation_date string,
last_update_time string)
USING delta location
's3://amzn-s3-demo-bucket/example-prefix/db/delta_table'""");

data.write.format("delta").mode("append").saveAsTable("delta_table")
```

SQL

```
-- Create a Delta  Lake table with the S3 location
CREATE TABLE delta_table(id string,
creation_date string,
last_update_time string)
USING delta LOCATION
's3://amzn-s3-demo-bucket/example-prefix/db/delta_table';

-- insert data into the table
INSERT INTO delta_table VALUES  ("100", "2015-01-01",
 "2015-01-01T13:51:39.340396Z"),
("101",  "2015-01-01",  "2015-01-01T12:14:58.597216Z"),
("102",  "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103",  "2015-01-01", "2015-01-01T13:51:40.519832Z");
```

Lecture à partir d'une table Delta Lake

PySpark

```
ddf = spark.table("delta_table")
ddf.show()
```

Scala

```
val ddf =  spark.table("delta_table")
ddf.show()
```

SQL

```
SELECT * FROM delta_table;
```

Utilisez un cluster Delta Lake avec Spark et AWS Glue

Pour utiliser le catalogue AWS Glue comme métastore pour les tables Delta Lake, créez un cluster en suivant les étapes suivantes. Pour plus d'informations sur la spécification de la classification de Delta Lake à l'aide de AWS Command Line Interface, voir Fournir une configuration à l'aide du AWS Command Line Interface lorsque vous créez un cluster ou Fournissez une configuration à l'aide du SDK Java lorsque vous créez un cluster.

Création d'un cluster Delta Lake

1. Créez un fichier, `configurations.json`, contenant les éléments suivants :

```
[{"Classification":"delta-defaults",
"Properties":{"delta.enabled":"true"}},
{"Classification":"spark-hive-site",
"Properties":
{"hive.metastore.client.factory.class":"com.amazonaws.glue.catalog.metastore.AWSGlueDataCat
```

2. Créez un cluster avec la configuration suivante, en remplaçant **`example Amazon S3 bucket path`** et **`subnet ID`** par les vôtres.

```
aws emr create-cluster
    --release-label  emr-6.9.0
    --applications Name=Spark
    --configurations file://delta_configurations.json
    --region us-east-1
    --name My_Spark_Delta_Cluster
    --log-uri  s3://amzn-s3-demo-bucket/
    --instance-type m5.xlarge
    --instance-count 2
    --service-role EMR_DefaultRole_V2
    --ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Considérations et restrictions

- Delta Lake est compatible avec les versions 6.9.0 et supérieures d'Amazon EMR. Vous pouvez utiliser Apache Spark 3.x sur des clusters Amazon EMR dotés de tables Delta.

- Nous vous recommandons d'utiliser le schéma d'URI s3 pour les chemins de localisation S3 au lieu de s3a pour des performances, une sécurité et une fiabilité optimales. Pour plus d'informations, consultez Utilisation des systèmes de stockage et de fichiers.
- Avec Amazon EMR 7.0, Delta Universal Format (UniForm) les convert-to-Iceberg instructions ne sont pas prises en charge.
- Avec Amazon EMR 6.9 et 6.10, lorsque vous stockez des données de table Delta Lake dans Amazon S3, les données de colonne deviennent `NULL` après l'opération de renommage des colonnes. Ce problème est résolu avec Amazon EMR 6.11. Pour plus d'informations sur l'opération expérimentale de changement de nom de colonne, consultez Opération de renommage de colonne dans le Guide de l'utilisateur de Delta Lake.
- Si vous créez une base de données dans le catalogue de données AWS Glue en dehors d'Apache Spark, le `LOCATION` champ de la base de données peut être vide. Comme Spark n'autorise pas la création de bases de données avec une propriété d'emplacement vide, l'erreur suivante s'affichera si vous utilisez Spark dans Amazon EMR pour créer une table Delta dans une base de données Glue et que la base de données possède une propriété vide `LOCATION` :

  ```
  IllegalArgumentException: Can not create a Path from an empty string
  ```

  Pour résoudre ce problème, créez la base de données dans le catalogue de données avec un chemin valide et non vide pour le champ `LOCATION`. Pour connaître les étapes de mise en œuvre de cette solution, consultez la section Exception d'argument illégal lors de la création d'une table dans le Guide de l'utilisateur Amazon Athena.

Historique des versions de Delta

Le tableau suivant répertorie les versions de Delta incluses dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version Delta

Étiquette de version Amazon EMR	Version Delta	Composants installés avec Delta
emr-7.9.0	3.3.0-amzn-1	Not available.

Étiquette de version Amazon EMR	Version Delta	Composants installés avec Delta
emr-7,8.0	3.3.0-amzn-0	Not available.
emr-7,7.0	3.2.1-amzn-2	Not available.
emr-7.6.0	3.2.1-amzn-1	Not available.
emr-7.5.0	3.2.1-amzn-0	Not available.
emr-7.4.0	3.2.0-amzn-1	Not available.
emr-7.3.0	3.2.0-amzn-0	Not available.
emr-7.2.0	3.1.0	Not available.
emr-7.1.0	3.0.0	Not available.
emr-7.0.0	3.0.0	Not available.
emr-6.15.0	2.4.0	Not available.
emr-6.14.0	2.4.0	Not available.
emr-6.13.0	2.4.0	Not available.
emr-6.12.0	2.4.0	Not available.
emr-6.11.1	2.2.0	Not available.
emr-6.11.0	2.2.0	Not available.
emr-6.10.1	2.2.0	Not available.
emr-6.10.0	2.2.0	Not available.
emr-6.9.1	2.1.0	Not available.
emr-6.9.0	2.1.0	Not available.

Apache Flink

Apache Flink est un moteur de flux de données en streaming permettant de traiter des flux en temps réel sur des sources de données à haut débit. Flink prend en charge la sémantique de l'heure des out-of-order événements pour les événements, la sémantique d'une seule fois, le contrôle de la contre-pression, et est APIs optimisé pour écrire à la fois des applications de streaming et des applications par lots.

En outre, Flink a des connecteurs pour les sources de données tierces, telles que :

- Amazon Kinesis Data Streams
- Apache Kafka
- Connecteur Flink Elasticsearch
- API de streaming Twitter
- Cassandra

Amazon EMR prend en charge Flink en tant qu'application YARN afin que vous puissiez gérer les ressources avec d'autres applications au sein d'un cluster. Flink-on-YARNvous permet de soumettre des tâches Flink transitoires, ou vous pouvez créer un cluster de longue durée qui accepte plusieurs tâches et alloue des ressources en fonction de la réservation YARN globale.

Flink est inclus dans les versions 5.1.0 et ultérieures d'Amazon EMR.

> **Note**
>
> La prise en charge de la classe `FlinkKinesisConsumer` a été ajoutée dans Amazon EMR version 5.2.1.

Le tableau suivant répertorie la version de Flink incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants qu'Amazon EMR installe avec Flink.

Pour la version des composants installés avec Flink dans cette version, consultez la version 7.9.0 Versions des composants.

Informations de version de Flink pour emr-7.9.0

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-7.9.0	Flink 1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Le tableau suivant répertorie la version de Flink incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Flink.

Pour connaître la version des composants installés avec Flink dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version de Flink pour emr-6.15.0

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-6.15.0	Flink 1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client,

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
		flink-jobmanager-config, hudi, delta-standalone-connectors

Le tableau suivant répertorie la version de Flink incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Flink.

Pour la version des composants installés avec Flink dans cette version, voir la version 5.36.2 Versions des composants.

Informations de version de Flink pour emr-5.36.2

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.36.2	Flink 1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Rubriques

- Création d'un cluster avec Flink
- Configuration de Flink dans Amazon EMR
- Utilisation des tâches Flink dans Amazon EMR
- Utilisation du shell Scala
- Recherche de l'interface web Flink
- Outil de mise à l’échelle automatique Flink

- Optimisation des temps de redémarrage pour les opérations de récupération et de mise à l'échelle de tâches
- Travailler avec les jobs Flink de Zeppelin dans Amazon EMR
- Historique des publications de Flink

Création d'un cluster avec Flink

Vous pouvez lancer un cluster avec le AWS Management Console AWS CLI, ou un AWS SDK.

Pour lancer un cluster avec Flink installé depuis la console

1. Ouvrez la console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com
2. Choisissez Créer un cluster et Go to advanced options (Aller aux options avancées).
3. Pour la Software Configuration (Configuration des logiciels), choisissez Version EMR emr-5.1.0 ou des versions ultérieures.
4. Choisissez Flink comme application, de même que toute autre application à installer.
5. Sélectionnez d'autres options si nécessaire, puis choisissez Create cluster (Créer un cluster).

Pour lancer un cluster avec Flink à partir de l' AWS CLI

- Créez le cluster à l'aide de la commande suivante :

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=MyKeyName,InstanceProfile=EMR_EC2_DefaultRole \
--steps Type=CUSTOM_JAR,Jar=command-runner.jar,Name=Flink_Long_Running_Session,\
Args=flink-yarn-session,-d
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

Configuration de Flink dans Amazon EMR

Configuration de Flink avec Hive Metastore et Glue Catalog

Les versions 6.9.0 et supérieures d'Amazon EMR prennent en charge à la fois Hive Metastore et AWS Glue Catalog avec le connecteur Apache Flink vers Hive. Cette section décrit les étapes nécessaires pour configurer AWS Glue Catalog et Hive Metastore avec Flink.

Rubriques

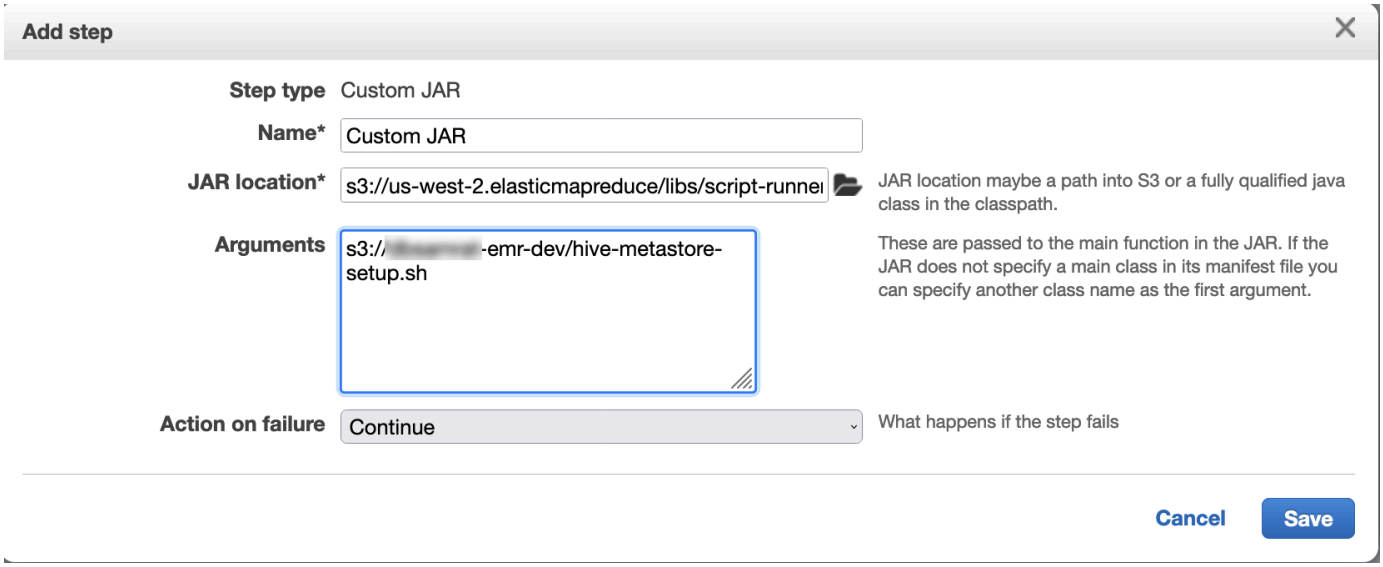
- Utilisation de Hive Metastore
- Utilisation du catalogue AWS de données Glue

Utilisation de Hive Metastore

1. Créez un cluster EMR avec la version 6.9.0 ou supérieure et au moins deux applications : Hive et Flink.
2. Utilisez script runner pour exécuter le script suivant en tant que fonction d'étape :

   `hive-metastore-setup.sh`

```
sudo cp /usr/lib/hive/lib/antlr-runtime-3.5.2.jar /usr/lib/flink/lib
sudo cp /usr/lib/hive/lib/hive-exec-3.1.3*.jar /lib/flink/lib
sudo cp /usr/lib/hive/lib/libfb303-0.9.3.jar /lib/flink/lib
sudo cp /usr/lib/flink/opt/flink-connector-hive_2.12-1.15.2.jar /lib/flink/lib
sudo chmod 755 /usr/lib/flink/lib/antlr-runtime-3.5.2.jar
sudo chmod 755 /usr/lib/flink/lib/hive-exec-3.1.3*.jar
sudo chmod 755 /usr/lib/flink/lib/libfb303-0.9.3.jar
sudo chmod 755 /usr/lib/flink/lib/flink-connector-hive_2.12-1.15.2.jar
```

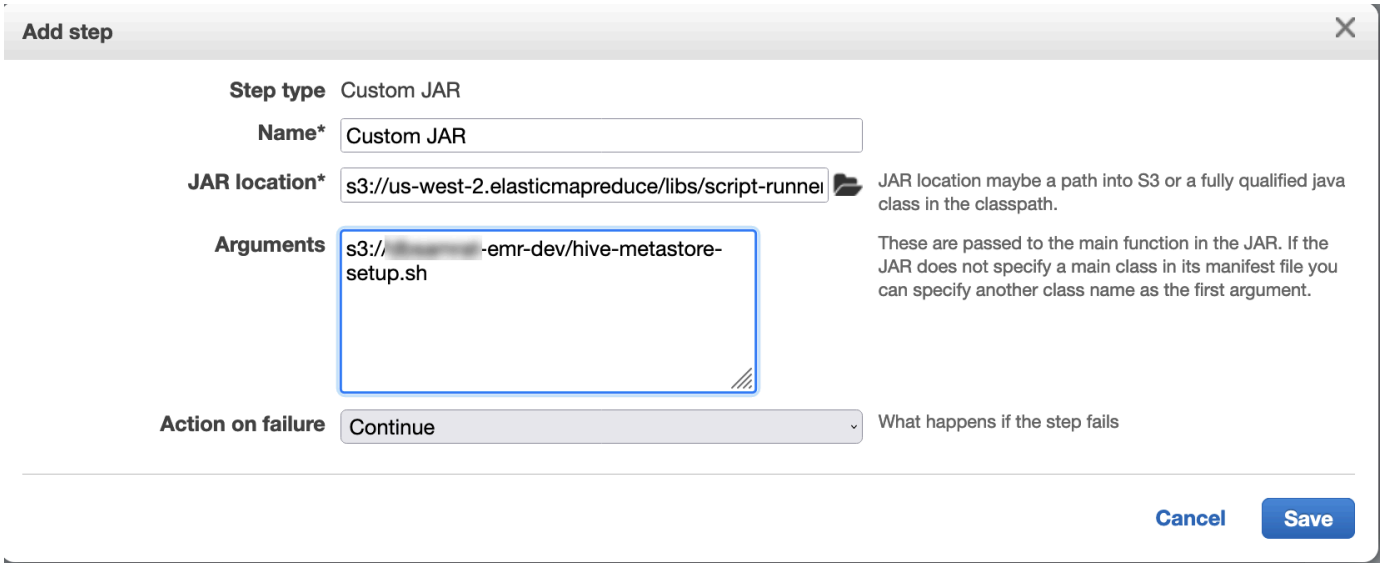


Utilisation du catalogue AWS de données Glue

1. Créez un cluster EMR avec la version 6.9.0 ou supérieure et au moins deux applications : Hive et Flink.
2. Sélectionnez Utiliser les métadonnées de la table Hive dans les paramètres du catalogue de données AWS Glue pour activer le catalogue de données dans le cluster.
3. Utilisez le script runner pour exécuter le script suivant en tant que fonction d'étape : Exécuter des commandes et des scripts sur un cluster Amazon EMR :

   glue-catalog-setup.sh

```
sudo cp /usr/lib/hive/auxlib/aws-glue-datacatalog-hive3-client.jar /usr/lib/flink/
lib
sudo cp /usr/lib/hive/lib/antlr-runtime-3.5.2.jar /usr/lib/flink/lib
sudo cp /usr/lib/hive/lib/hive-exec-3.1.3*.jar /lib/flink/lib
sudo cp /usr/lib/hive/lib/libfb303-0.9.3.jar /lib/flink/lib
sudo cp /usr/lib/flink/opt/flink-connector-hive_2.12-1.15.2.jar /lib/flink/lib
sudo chmod 755 /usr/lib/flink/lib/aws-glue-datacatalog-hive3-client.jar
sudo chmod 755 /usr/lib/flink/lib/antlr-runtime-3.5.2.jar
sudo chmod 755 /usr/lib/flink/lib/hive-exec-3.1.3*.jar
sudo chmod 755 /usr/lib/flink/lib/libfb303-0.9.3.jar
sudo chmod 755 /usr/lib/flink/lib/flink-connector-hive_2.12-1.15.2.jar
```



Configurer Flink avec un fichier de configuration

Vous pouvez utiliser l'API de configuration Amazon EMR pour configurer Flink à l'aide d'un fichier de configuration. Les fichiers configurables dans l'API sont les suivants :

- `flink-conf.yaml`
- `log4j.properties`
- `flink-log4j-session`
- `log4j-cli.properties`

Le fichier de configuration principal de Flink est `flink-conf.yaml`.

Pour configurer le nombre d'emplacements de tâches utilisés par Flink à partir du fichier AWS CLI

1. Créez un fichier, `configurations.json`, contenant les éléments suivants :

```
[
    {
      "Classification": "flink-conf",
      "Properties": {
        "taskmanager.numberOfTaskSlots":"2"
      }
    }
```

```
]
```

2. Créez ensuite un cluster à l'aide de la configuration suivante :

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=YourKeyName,InstanceProfile=EMR_EC2_DefaultRole
```

> **Note**
>
> Vous pouvez également modifier certaines configurations avec l'API Flink. Pour plus d'informations, consultez Concepts dans la documentation Flink.
> Avec la version 5.21.0 et ultérieures d'Amazon EMR, vous permet de remplacer les configurations de cluster et de spécifier des classifications de configuration supplémentaires pour chaque groupe d'instances dans un cluster en cours d'exécution. Pour ce faire, utilisez la console Amazon EMR, le AWS Command Line Interface (AWS CLI) ou le AWS SDK. Pour plus d'informations, consultez Fourniture d'une configuration pour un groupe d'instances dans un cluster en cours d'exécution.

Options de parallélisme

En tant que propriétaire de votre application, c'est vous qui savez le mieux quelles ressources attribuer aux tâches dans Flink. Pour les exemples de cette documentation, utilisez le même nombre de tâches que les instances de tâches que vous utilisez pour l'application. Généralement, nous recommandons cela pour le niveau initial de parallélisme, mais vous pouvez également augmenter la granularité de parallélisme à l'aide des emplacements de tâches, qui ne doivent généralement pas dépasser le nombre de cœurs virtuels par instance. Pour plus d'informations sur l'architecture de Flink, consultez Concepts dans la documentation Flink.

Configuration de Flink sur un cluster EMR doté de plusieurs nœuds primaires

Flink reste disponible pendant le processus JobManager de basculement du nœud primaire dans un cluster Amazon EMR comportant plusieurs nœuds primaires. À partir d'Amazon EMR 5.28.0, la JobManager haute disponibilité est également activée automatiquement. Aucune configuration manuelle n'est nécessaire.

Avec Amazon EMR versions 5.27.0 ou antérieures, il s' JobManager agit d'un point de défaillance unique. En cas d' JobManager échec, il perd tous les états des tâches et ne reprend pas les tâches en cours d'exécution. Vous pouvez activer la JobManager haute disponibilité en configurant le nombre de tentatives d'application, le point de contrôle et ZooKeeper en activant le stockage d'état pour Flink, comme le montre l'exemple suivant :

```
[
  {
    "Classification": "yarn-site",
    "Properties": {
      "yarn.resourcemanager.am.max-attempts": "10"
    }
  },
  {
    "Classification": "flink-conf",
    "Properties": {
        "yarn.application-attempts": "10",
        "high-availability": "zookeeper",
        "high-availability.zookeeper.quorum": "%{hiera('hadoop::zk')}",
        "high-availability.storageDir": "hdfs:///user/flink/recovery",
        "high-availability.zookeeper.path.root": "/flink"
    }
  }
]
```

Vous devez configurer le nombre maximal de tentatives du maître d'application pour YARN et les tentatives d'application pour Flink. Pour plus d'informations, consultez Configuration de la haute disponibilité du cluster YARN. Vous pouvez également configurer le point de contrôle Flink pour que les tâches de JobManager restauration en cours d'exécution redémarrées soient restaurées à partir de points de contrôle précédemment terminés. Pour plus d'informations, consultez Flink checkpointing.

Configuration de la taille du processus de mémoire

Pour les versions d'Amazon EMR qui utilisent Flink 1.11.x, vous devez configurer la taille totale du processus de mémoire pour () et JobManager (`jobmanager.memory.process.size`) dans. TaskManager `taskmanager.memory.process.size` `flink-conf.yaml` Vous pouvez définir ces valeurs soit en configurant le cluster à l'aide de l'API de configuration, soit en décommentant manuellement ces champs via SSH. Flink fournit les valeurs par défaut suivantes.

- `jobmanager.memory.process.size` : 1600m
- `taskmanager.memory.process.size` : 1728m

Pour exclure le métaspace JVM et la surcharge, utilisez la taille de mémoire totale de Flink (`taskmanager.memory.flink.size`) au lieu de `taskmanager.memory.process.size`. La valeur par défaut du paramètre `taskmanager.memory.process.size` est 1280m. Il n'est pas recommandé de définir à la fois `taskmanager.memory.process.size` et `taskmanager.memory.process.size`.

Toutes les versions d'Amazon EMR qui utilisent Flink 1.12.0 et versions ultérieures ont les valeurs par défaut répertoriées dans l'ensemble open source pour Flink comme valeurs par défaut sur Amazon EMR. Vous n'avez donc pas besoin de les configurer vous-même.

Configuration de la taille du fichier de sortie du journal

Les conteneurs d'applications Flink créent et écrivent dans trois types de fichiers journaux : fichiers `.out`, fichiers `.log` et fichiers `.err`. Seuls `.err` les fichiers sont compressés et supprimés du système de fichiers, tandis que les fichiers journaux `.log` et `.out` restent dans le système de fichiers. Pour garantir la gérabilité de ces fichiers de sortie et la stabilité du cluster, vous pouvez configurer la rotation des journaux dans `log4j.properties` afin de définir un nombre maximum de fichiers et de limiter leur taille.

Amazon EMR versions 5.30.0 et ultérieures

À partir d'Amazon EMR 5.30.0, Flink utilise le framework de journalisation log4j2 avec le nom de classification de configuration `flink-log4j.`. L'exemple de configuration suivant illustre le format log4j2.

```
[
  {
```

```
    "Classification": "flink-log4j",
    "Properties": {
      "appender.main.name": "MainAppender",
      "appender.main.type": "RollingFile",
      "appender.main.append" : "false",
      "appender.main.fileName" : "${sys:log.file}",
      "appender.main.filePattern" : "${sys:log.file}.%i",
      "appender.main.layout.type" : "PatternLayout",
      "appender.main.layout.pattern" : "%d{yyyy-MM-dd HH:mm:ss,SSS} %-5p %-60c %x - %m
%n",
      "appender.main.policies.type" : "Policies",
      "appender.main.policies.size.type" : "SizeBasedTriggeringPolicy",
      "appender.main.policies.size.size" : "100MB",
      "appender.main.strategy.type" : "DefaultRolloverStrategy",
      "appender.main.strategy.max" : "10"
    },
  }
]
```

Amazon EMR versions 5.29.0 et antérieures

Avec les versions 5.29.0 et antérieures d'Amazon EMR, Flink utilise le framework de journalisation log4j. L'exemple de configuration suivant illustre le format log4j.

```
[
  {
    "Classification": "flink-log4j",
    "Properties": {
      "log4j.appender.file": "org.apache.log4j.RollingFileAppender",
      "log4j.appender.file.append":"true",
      # keep up to 4 files and each file size is limited to 100MB
      "log4j.appender.file.MaxFileSize":"100MB",
      "log4j.appender.file.MaxBackupIndex":4,
      "log4j.appender.file.layout":"org.apache.log4j.PatternLayout",
      "log4j.appender.file.layout.ConversionPattern":"%d{yyyy-MM-dd HH:mm:ss,SSS} %-5p
 %-60c %x - %m%n"
    },
  }
]
```

Configurer Flink pour qu'il fonctionne avec Java 11

Les versions 6.12.0 et supérieures d'Amazon EMR fournissent un support d'exécution Java 11 pour Flink. Les sections suivantes décrivent comment configurer le cluster pour fournir un support d'exécution Java 11 pour Flink.

Rubriques

- Configurer Flink pour Java 11 lorsque vous créez un cluster
- Configurer Flink pour Java 11 sur un cluster en cours d'exécution
- Confirmez le runtime Java pour Flink sur un cluster en cours d'exécution

Configurer Flink pour Java 11 lorsque vous créez un cluster

Procédez comme suit pour créer un cluster EMR avec Flink et Java 11 Runtime. Le fichier de configuration dans lequel vous ajoutez la prise en charge de Java 11 est `flink-conf.yaml`.

Console

Pour créer un cluster avec Flink et Java 11 Runtime dans la console

1. Connectez-vous à la et ouvrez AWS Management Console la console Amazon EMR à https://console.aws.amazon.com l'adresse /emr.
2. Choisissez Clusters sous EMR activé EC2 dans le volet de navigation, puis Create cluster.
3. Sélectionnez Amazon EMR version 6.12.0 ou supérieure, puis choisissez d'installer l'application Flink. Sélectionnez les autres applications que vous souhaitez installer sur votre cluster.
4. Poursuivez la configuration de votre cluster. Dans la section facultatifs Paramètres logiciels, utilisez l'option par défaut Entrer un configuration et entrez la configuration suivante :

```
[
    {
      "Classification": "flink-conf",
      "Properties": {
        "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "env.java.home":"/usr/lib/jvm/jre-11"
      }
    }
```

```
]
```

5. Continuez à configurer et à lancer votre cluster.

AWS CLI

Pour créer un cluster avec Flink et Java 11 Runtime à partir de la CLI

1. Créez un fichier de configuration `configurations.json` qui configure Flink pour utiliser Java 11.

```
[
    {
      "Classification": "flink-conf",
      "Properties": {
        "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "env.java.home":"/usr/lib/jvm/jre-11"
      }
    }
]
```

2. À partir de AWS CLI, créez un nouveau cluster EMR avec Amazon EMR version 6.12.0 ou supérieure, et installez l'application Flink, comme indiqué dans l'exemple suivant :

```
aws emr create-cluster --release-label emr-6.12.0 \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=YourKeyName,InstanceProfile=EMR_EC2_DefaultRole
```

Configurer Flink pour Java 11 sur un cluster en cours d'exécution

Procédez comme suit pour mettre à jour un cluster EMR en cours d'exécution avec Flink et Java 11 Runtime. Le fichier de configuration dans lequel vous ajoutez la prise en charge de Java 11 est `flink-conf.yaml`.

Console

Pour mettre à jour un cluster en cours d'exécution avec Flink et Java 11 Runtime dans la console

1. Connectez-vous à la et ouvrez AWS Management Console la console Amazon EMR à https://console.aws.amazon.com l'adresse /emr.
2. Sélectionnez Clusters sous EMR activé EC2 dans le volet de navigation, puis sélectionnez le cluster que vous souhaitez mettre à jour.

   > **Note**
   >
   > Le cluster doit utiliser Amazon EMR version 6.12.0 ou ultérieure pour prendre en charge Java 11.

3. Sélectionnez l'onglet Configurations.
4. Dans la section Configurations du groupe d'instances, sélectionnez le groupe d'instances en cours d'exécution que vous souhaitez mettre à jour, puis choisissez Reconfigurer dans le menu d'actions de la liste.
5. Reconfigurez le groupe d'instances avec l'option Modifier les attributs comme suit. Sélectionnez Ajouter une nouvelle configuration après chacune d'entre elles.

Classification	Propriété	Valeur
`flink-conf`	`containerized.taskmanager.env.JAVA_HOME`	`/usr/lib/jvm/jre-11`
`flink-conf`	`containerized.master.env.JAVA_HOME`	`/usr/lib/jvm/jre-11`
`flink-conf`	`env.java.home`	`/usr/lib/jvm/jre-11`

6. Sélectionnez Enregistrer les modifications pour ajouter les configurations.

AWS CLI

Pour mettre à jour un cluster en cours d'exécution afin d'utiliser Flink et le runtime Java 11 à partir de la CLI

Utilisez la commande `modify-instance-groups` pour spécifier une nouvelle configuration pour un groupe d'instances dans un cluster en cours d'exécution.

1. Créez d'abord un fichier de configuration `configurations.json` qui configure Flink pour utiliser Java 11. Dans l'exemple suivant, remplacez *ig-1xxxxxxx9* par l'ID du groupe d'instances que vous souhaitez reconfigurer. Enregistrez le fichier dans le même répertoire que celui où vous exécuterez la commande `modify-instance-groups`.

```
[
   {
      "InstanceGroupId":"ig-1xxxxxxx9",
      "Configurations":[
         {
            "Classification":"flink-conf",
            "Properties":{
              "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
              "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
              "env.java.home":"/usr/lib/jvm/jre-11"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. Dans l' AWS CLI, exécutez la commande suivante. Remplacez l'ID du groupe d'instances que vous souhaitez reconfigurer :

```
aws emr modify-instance-groups --cluster-id j-2AL4XXXXXX5T9 \
--instance-groups file://configurations.json
```

Confirmez le runtime Java pour Flink sur un cluster en cours d'exécution

Pour déterminer l'environnement d'exécution Java d'un cluster en cours d'exécution, connectez-vous au nœud primaire avec SSH, comme décrit dans Connexion au nœud primaire avec SSH. Ensuite, exécutez la commande suivante :

```
ps -ef | grep flink
```

La commande `ps` associée à l'option `-ef` répertorie tous les processus en cours d'exécution sur le système. Vous pouvez filtrer cette sortie avec `grep` pour trouver les mentions de la chaîne `flink`. Vérifiez le résultat pour la valeur de l'environnement d'exécution Java (JRE), `jre-XX`. Dans le résultat suivant, `jre-11` indique que Java 11 est activé lors de l'exécution de Flink.

```
flink    19130     1  0 09:17 ?        00:00:15 /usr/lib/jvm/jre-11/bin/
java -Djava.io.tmpdir=/mnt/tmp -Dlog.file=/usr/lib/flink/log/flink-flink-
historyserver-0-ip-172-31-32-127.log -Dlog4j.configuration=file:/usr/lib/flink/conf/
log4j.properties -Dlog4j.configurationFile=file:/usr/lib/flink/conf/log4j.properties
 -Dlogback.configurationFile=file:/usr/lib/flink/conf/logback.xml -classpath /usr/lib/
flink/lib/flink-cep-1.17.0.jar:/usr/lib/flink/lib/flink-connector-files-1.17.0.jar:/
usr/lib/flink/lib/flink-csv-1.17.0.jar:/usr/lib/flink/lib/flink-json-1.17.0.jar:/usr/
lib/flink/lib/flink-scala_2.12-1.17.0.jar:/usr/lib/flink/lib/flink-table-api-java-
uber-1.17.0.jar:/usr/lib/flink/lib/flink-table-api-scala-bridge_2.12-1.17.0.
```

Vous pouvez également vous connecter au nœud primaire avec SSH et démarrer une session Flink YARN avec une commande `flink-yarn-session -d`. La sortie montre la machine virtuelle Java (JVM) pour Flink, `java-11-amazon-corretto` dans l'exemple suivant :

```
2023-05-29 10:38:14,129 INFO  org.apache.flink.configuration.GlobalConfiguration
     [] - Loading configuration property: containerized.master.env.JAVA_HOME, /usr/lib/
jvm/java-11-amazon-corretto.x86_64
```

Utilisation des tâches Flink dans Amazon EMR

Il existe plusieurs manières d'interagir avec Flink sur Amazon EMR : via la console, via l'interface Flink située dans l'interface utilisateur de suivi et via ResourceManager la ligne de commande. Vous pouvez envoyer un fichier JAR à une application Flink avec n'importe lequel de ces outils. Une fois soumis un fichier JAR, il devient une tâche gérée par le Flink JobManager. JobManager Il se trouve sur le nœud YARN qui héberge le daemon Application Master de la session Flink.

Vous pouvez exécuter une application Flink en tant que tâche YARN sur un cluster à exécution longue ou sur un cluster transitoire. Sur un cluster à exécution longue, vous pouvez soumettre plusieurs tâches Flink à un cluster Flink s'exécutant sur Amazon EMR. Si vous exécutez une tâche Flink sur un cluster transitoire, votre cluster Amazon EMR n'existe que le temps de l'exécution de l'application Flink, et vous n'êtes donc facturé que pour les ressources et le temps utilisés. Vous pouvez soumettre une tâche Flink avec l'opération d'API Amazon `AddSteps` EMR, en tant qu'argument d'étape de `RunJobFlow` l'opération et via AWS CLI `add-steps` les `create-cluster` commandes or.

Démarrez une application Flink YARN en tant qu'étape sur un cluster de longue durée

Pour démarrer une application Flink à laquelle plusieurs clients peuvent soumettre du travail via les opérations de l'API YARN, vous devez soit créer un cluster, soit ajouter une application Flink à un cluster existant. Pour obtenir des instructions sur la création d'un cluster, veuillez consulter Création d'un cluster avec Flink. Pour démarrer une session YARN sur un cluster existant, procédez comme suit à partir de la console, de l' AWS CLI ou du kit SDK Java.

> **Note**
>
> La commande `flink-yarn-session` a été ajoutée dans la version 5.5.0 d'Amazon EMR comme habillage pour le script `yarn-session.sh`, afin de simplifier l'exécution. Si vous utilisez une version antérieure d'Amazon EMR, remplacez `bash -c "/usr/lib/flink/bin/yarn-session.sh -d"` par Arguments dans la console ou par `Args`. dans la commande AWS CLI .

Pour soumettre une tâche Flink sur un cluster existant depuis la console

Soumettez la session Flink avec la commande `flink-yarn-session` dans un cluster existant.

1. Ouvrez la console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com
2. Dans la liste des clusters, sélectionnez celui que vous avez lancé précédemment.
3. Dans la page des détails de cluster, sélectionnez Steps (Étapes), Add Step (Ajouter une étape).
4. Utilisez les directives suivantes pour saisir les paramètres, puis sélectionnez Ajouter.

Paramètre	Description
Type d'étape	JAR personnalisé
Nom	Nom permettant d'identifier la nouvelle étape. Par exemple, *<example-flink-step-name>*.
Emplacement JAR	command-runner.jar
Arguments	La commande `flink-yarn-session` avec les arguments appropriés pour votre application. Par exemple, *`flink-yarn-session -d`* démarre une session Flink au sein de votre cluster YARN dans un état détaché (`-d`). Consultez Configuration de YARN dans la dernière documentation Flink pour plus de détails sur les arguments.

Pour soumettre une tâche Flink sur un cluster existant à l'aide du AWS CLI

- Utilisez la commande `add-steps` pour ajouter une tâche Flink à un cluster de longue durée. L'exemple de commande suivant indique `Args="flink-yarn-session", "-d"` pour démarrer une session Flink au sein de votre cluster YARN dans un état détaché (`-d`). Consultez Configuration de YARN dans la dernière documentation Flink pour plus de détails sur les arguments.

```
aws emr add-steps --cluster-id <j-XXXXXXXX> --steps Type=CUSTOM_JAR,Name=<example-flink-step-name>,Jar=command-runner.jar,Args="flink-yarn-session","-d"
```

Soumettre du travail à une application Flink existante sur un cluster de longue durée

Si vous possédez déjà une application Flink sur un cluster de longue durée, vous pouvez spécifier l'ID de l'application Flink du cluster afin d'y soumettre du travail. Pour obtenir l'ID de l'application, `yarn application -list` exécutez l'opération YarnClientAPI AWS CLI ou via l'API :

```
$ yarn application -list
16/09/07 19:32:13 INFO client.RMProxy: Connecting to ResourceManager at
 ip-10-181-83-19.ec2.internal/10.181.83.19:8032
Total number of applications (application-types: [] and states: [SUBMITTED, ACCEPTED,
 RUNNING]):1
Application-Id    Application-Name    Application-Type    User    Queue    State
 Final-State    Progress    Tracking-URL
application_1473169569237_0002    Flink session with 14 TaskManagers (detached)
   Apache Flink     hadoop    default            RUNNING          UNDEFINED
 100% http://ip-10-136-154-194.ec2.internal:33089
```

L'identifiant de l'application pour cette session Flink est`application_1473169569237_0002`, que vous pouvez utiliser pour soumettre du travail à l'application à partir du SDK AWS CLI ou d'un SDK.

Example SDK pour Java

```
List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

HadoopJarStepConfig flinkWordCountConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("flink", "run", "-m", "yarn-cluster", "-yid",
 "application_1473169569237_0002", "-yn", "2", "/usr/lib/flink/examples/streaming/
WordCount.jar",
      "--input", "s3://amzn-s3-demo-bucket/pg11.txt", "--output", "s3://amzn-s3-demo-
bucket/alice2/");

StepConfig flinkRunWordCount = new StepConfig()
  .withName("Flink add a wordcount step")
  .withActionOnFailure("CONTINUE")
  .withHadoopJarStep(flinkWordCountConf);

stepConfigs.add(flinkRunWordCount);

AddJobFlowStepsResult res = emr.addJobFlowSteps(new AddJobFlowStepsRequest()
   .withJobFlowId("myClusterId")
```

```
    .withSteps(stepConfigs));
```

Example AWS CLI

```
aws emr add-steps --cluster-id <j-XXXXXXXX> \
--steps Type=CUSTOM_JAR,Name=Flink_Submit_To_Long_Running,Jar=command-runner.jar,\
Args="flink","run","-m","yarn-cluster","-yid","application_1473169569237_0002",\
"/usr/lib/flink/examples/streaming/WordCount.jar",\
"--input","s3://amzn-s3-demo-bucket/pg11.txt","--output","s3://amzn-s3-demo-bucket/
alice2/" \
--region <region-code>
```

Soumission d'une tâche Flink transitoire

Les exemples suivants lancent un cluster transitoire qui exécute une tâche Flink, puis la résilie à son terme.

Example SDK pour Java

```
import java.util.ArrayList;
import java.util.List;
import com.amazonaws.AmazonClientException;
import com.amazonaws.auth.AWSCredentials;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.profile.ProfileCredentialsProvider;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduce;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduceClientBuilder;
import com.amazonaws.services.elasticmapreduce.model.*;

public class Main_test {

 public static void main(String[] args) {
  AWSCredentials credentials_profile = null;
  try {
   credentials_profile = new ProfileCredentialsProvider("default").getCredentials();
  } catch (Exception e) {
   throw new AmazonClientException(
     "Cannot load credentials from .aws/credentials file. " +
       "Make sure that the credentials file exists and the profile name is specified
 within it.",
     e);
  }
```

```
  AmazonElasticMapReduce emr = AmazonElasticMapReduceClientBuilder.standard()
    .withCredentials(new AWSStaticCredentialsProvider(credentials_profile))
    .withRegion(Regions.US_WEST_1)
    .build();

  List<StepConfig> stepConfigs = new ArrayList<StepConfig>();
  HadoopJarStepConfig flinkWordCountConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("bash", "-c", "flink", "run", "-m", "yarn-cluster", "-yn", "2",
      "/usr/lib/flink/examples/streaming/WordCount.jar", "--input", "s3://path/to/
input-file.txt", "--output",
      "s3://path/to/output/");

  StepConfig flinkRunWordCountStep = new StepConfig()
    .withName("Flink add a wordcount step and terminate")
    .withActionOnFailure("CONTINUE")
    .withHadoopJarStep(flinkWordCountConf);

  stepConfigs.add(flinkRunWordCountStep);

  Application flink = new Application().withName("Flink");

  RunJobFlowRequest request = new RunJobFlowRequest()
    .withName("flink-transient")
    .withReleaseLabel("emr-5.20.0")
    .withApplications(flink)
    .withServiceRole("EMR_DefaultRole")
    .withJobFlowRole("EMR_EC2_DefaultRole")
    .withLogUri("s3://path/to/my/logfiles")
    .withInstances(new JobFlowInstancesConfig()
      .withEc2KeyName("myEc2Key")
      .withEc2SubnetId("subnet-12ab3c45")
      .withInstanceCount(3)
      .withKeepJobFlowAliveWhenNoSteps(false)
      .withMasterInstanceType("m4.large")
      .withSlaveInstanceType("m4.large"))
    .withSteps(stepConfigs);

  RunJobFlowResult result = emr.runJobFlow(request);
  System.out.println("The cluster ID is " + result.toString());

 }
```

```
}
```

Example AWS CLI

Utilisez la sous-commande `create-cluster` pour créer un cluster transitoire qui se termine quand la tâche Flink est terminée :

```
aws emr create-cluster --release-label emr-5.2.1 \
--name "Flink_Transient" \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--auto-terminate
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=<YourKeyName>,InstanceProfile=EMR_EC2_DefaultRole \
--steps Type=CUSTOM_JAR,Jar=command-runner.jar,Name=Flink_Long_Running_Session,\
Args="bash","-c","\"flink run -m yarn-cluster /usr/lib/flink/examples/streaming/
WordCount.jar
--input s3://amzn-s3-demo-bucket/pg11.txt --output s3://amzn-s3-demo-bucket/alice/""
```

Utilisation du shell Scala

Le shell Scala Flink pour les clusters EMR n'est configuré que pour démarrer de nouvelles sessions de YARN. Vous pouvez utiliser le shell Scala en suivant la procédure ci-dessous.

Utilisation du shell Flink Scala sur le nœud primaire

1. Connectez-vous au nœud primaire avec SSH comme décrit dans Connexion au nœud primaire avec SSH.
2. Tapez ce qui suit pour démarrer un shell :

   Dans Amazon EMR version 5.5.0 et versions ultérieures, vous pouvez utiliser la commande suivante pour démarrer un cluster Yarn pour le Scala Shell avec un. TaskManager

   ```
   % flink-scala-shell yarn 1
   ```

   Dans les versions antérieures d'Amazon EMR, utilisez :

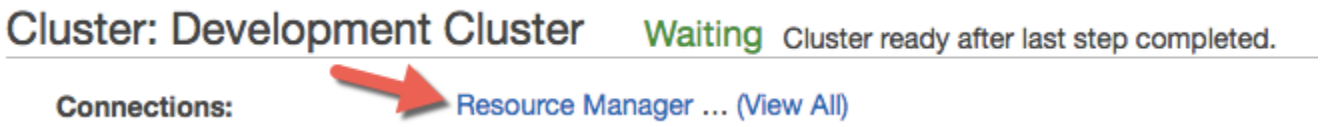
```
% /usr/lib/flink/bin/start-scala-shell.sh yarn 1
```

Cela démarre le shell Scala Flink pour que vous puissiez utiliser Flink de façon interactive. Tout comme avec les autres interfaces et options, vous pouvez faire évoluer la valeur de l'option `-n` utilisée dans l'exemple basé sur le nombre de tâches à exécuter à partir du shell.

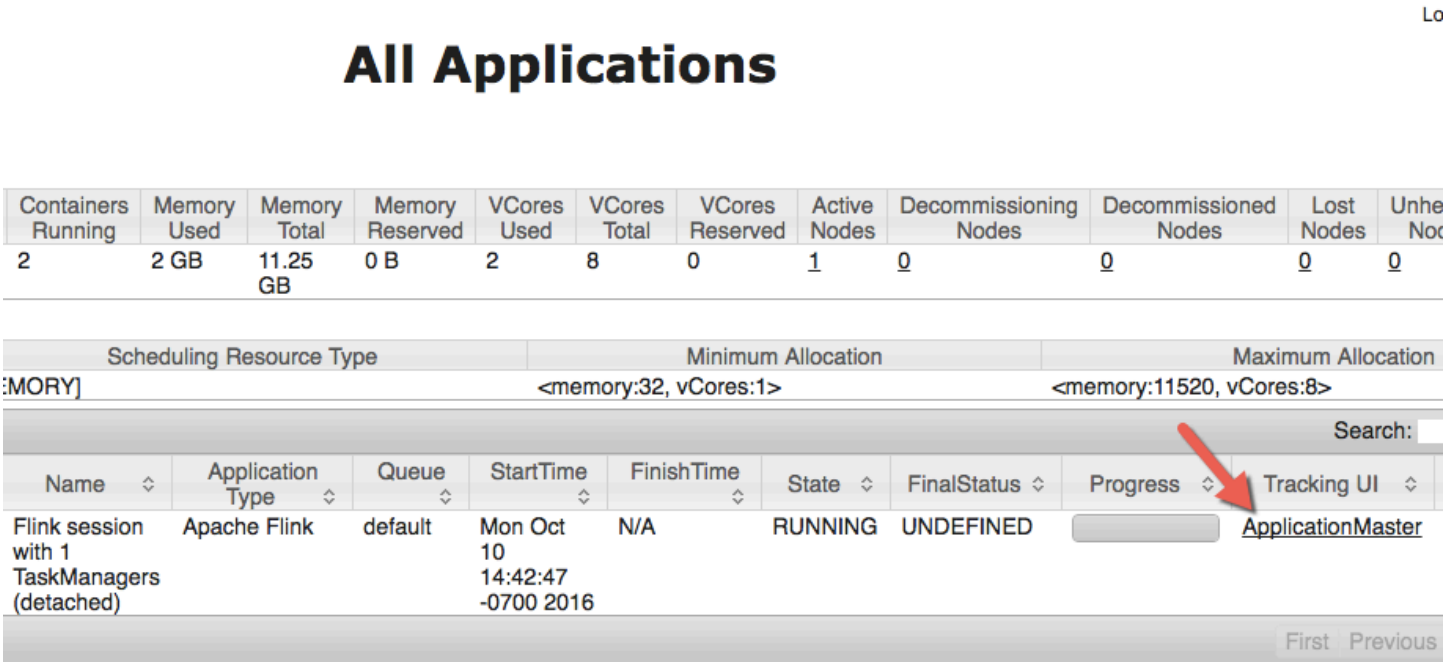
Pour plus d'informations, consultez Scala REPL dans la documentation officielle d'Apache Flink.

Recherche de l'interface web Flink

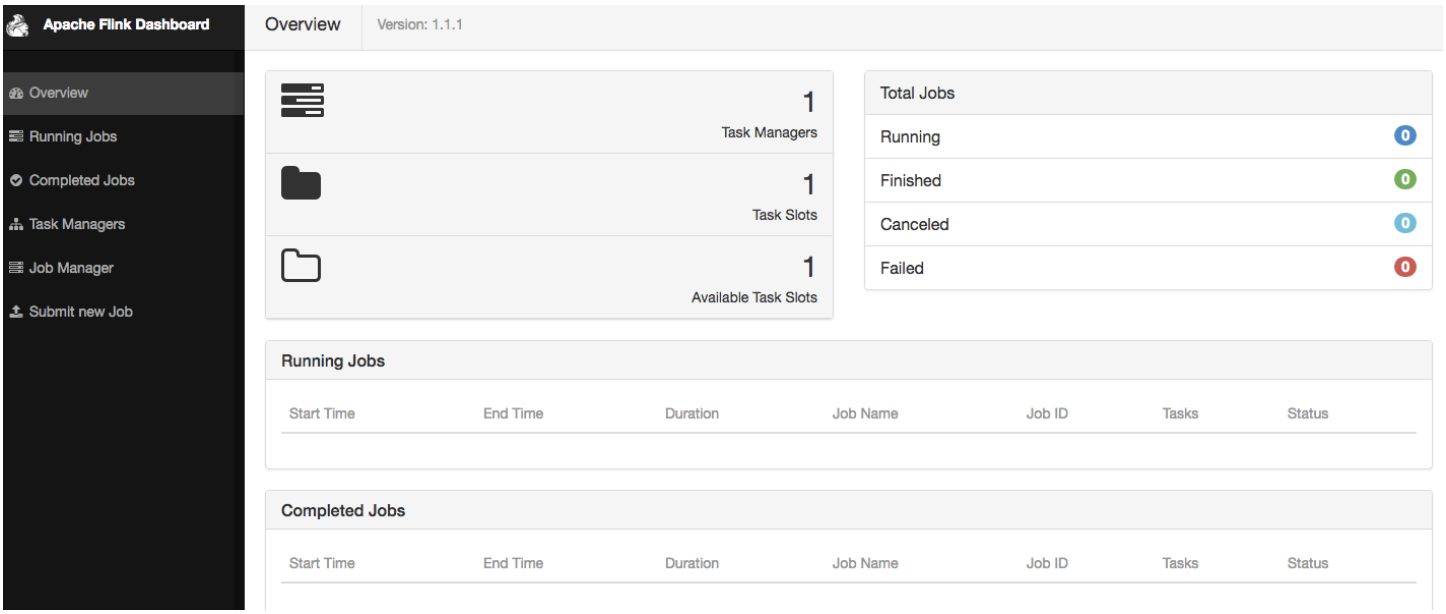
Le maître d'application qui appartient à l'application Flink héberge l'interface Web de Flink. Il s'agit d'une autre méthode pour soumettre un fichier JAR en tant que tâche ou pour afficher le statut actuel d'autres tâches. L'interface Web Flink est active tant qu'une session Flink est en cours d'exécution. Si une tâche YARN de longue date est déjà active, vous pouvez suivre les instructions de la rubrique Connexion au nœud principal avec SSH du guide de gestion Amazon EMR pour vous connecter au YARN. ResourceManager Par exemple, si vous avez configuré un tunnel SSH et que vous avez activé un proxy dans votre navigateur, vous choisissez la ResourceManager connexion sous Connexions sur la page de détails de votre cluster EMR.



Une fois que vous avez trouvé le ResourceManager, sélectionnez l'application YARN qui héberge une session Flink. Cliquez sur le lien sous la colonne Tracking UI (UI de suivi).



Dans l'interface Web Flink, vous pouvez afficher la configuration, soumettre votre propre JAR personnalisé comme une tâche, ou contrôler les travaux en cours.



Outil de mise à l'échelle automatique Flink

Présentation

La version 6.15.0 et les versions ultérieures d'Amazon EMR prennent en charge l'outil de mise à l'échelle automatique Flink. La fonctionnalité de mise à l'échelle automatique des tâches collecte les

métriques issues de l'exécution des tâches de streaming Flink et met à l'échelle automatiquement les vertex de chaque tâche. Cela réduit la contre-pression et répond à l'objectif d'utilisation que vous avez défini.

Pour plus d'informations, voir la rubrique Outil de mise à l'échelle automatique de la documentation de l'opérateur Apache Flink Kubernetes.

Considérations

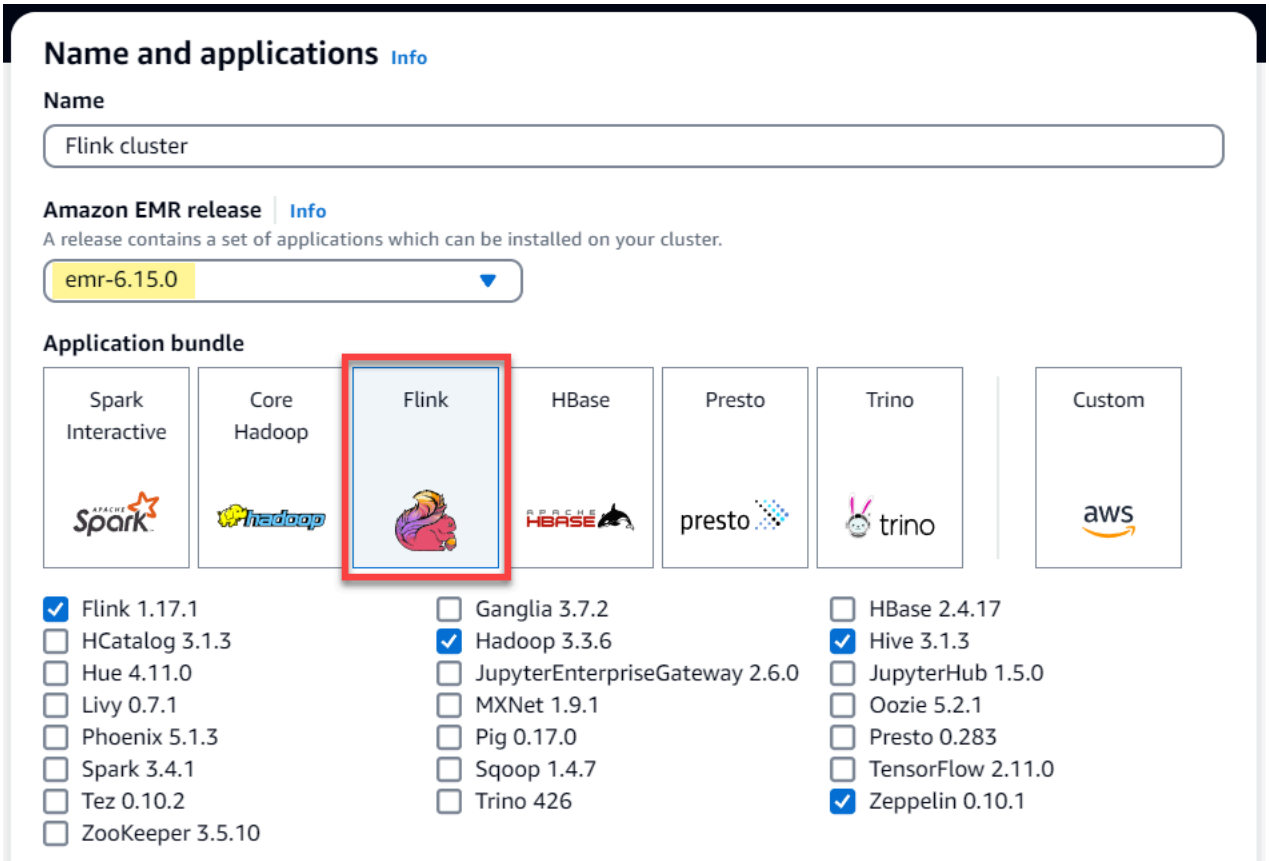
- La version 6.15.0 et les versions ultérieures d'Amazon EMR prennent en charge l'outil de mise à l'échelle automatique Flink.
- L'outil de mise à l'échelle automatique Flink est pris en charge uniquement pour les tâches de streaming.
- Seul le planificateur adaptatif est pris en charge. Le planificateur par défaut n'est pas pris en charge.
- Nous recommandons d'activer la mise à l'échelle de cluster pour permettre une mise à disposition de ressources dynamique. La mise à l'échelle gérée par Amazon EMR est préférable, car les métriques sont évaluées toutes les 5 à 10 secondes. À cet intervalle, votre cluster peut s'adapter plus facilement à l'évolution des ressources de cluster requises.

Activation de l'outil de mise à l'échelle automatique

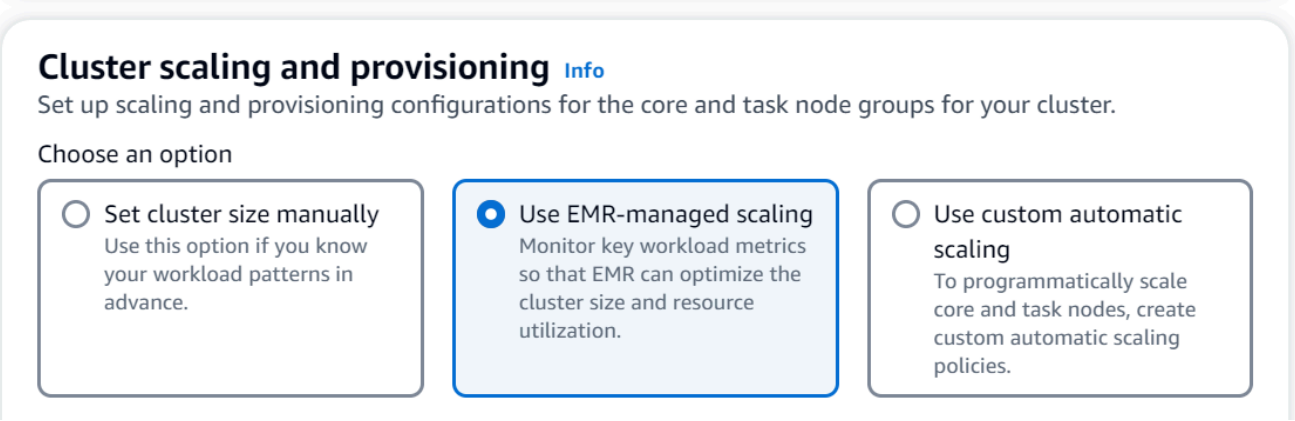
Suivez les étapes ci-dessous pour activer l'autoscaler Flink lorsque vous créez un Amazon EMR sur un cluster. EC2

1. Créez un cluster EMR depuis la console Amazon EMR :

    a. Choisissez Amazon EMR version `emr-6.15.0` ou ultérieure. Sélectionnez l'offre d'applications Flink, puis toutes les autres applications que vous souhaitez inclure dans votre cluster.



b. Dans Mise à l'échelle et provisionnement du cluster, sélectionnez l'option Utiliser la mise à l'échelle gérée par EMR.



2. Dans la section Paramètres du logiciel, entrez la configuration suivante pour activer l'outil de mise à l'échelle automatique Flink. Pour les scénarios de test, définissez l'intervalle de décision, l'intervalle entre les fenêtres de mesures et l'intervalle de stabilisation sur une valeur inférieure afin que la tâche prenne immédiatement une décision de mise à l'échelle et faciliter ainsi la vérification.

```
[
```

```
  {
    "Classification": "flink-conf",
    "Properties": {
      "job.autoscaler.enabled": "true",
      "jobmanager.scheduler": "adaptive",
      "job.autoscaler.stabilization.interval": "60s",
      "job.autoscaler.metrics.window": "60s",
      "job.autoscaler.decision.interval": "10s",
      "job.autoscaler.debug.logs.interval": "60s"
    }
  }
]
```

3. Sélectionnez ou configurez les autres paramètres selon vos préférences, puis créez le cluster compatible avec l'outil de mise à l'échelle automatique Flink.

Configurations de l'outil de mise à l'échelle automatique

Cette section aborde la plupart des configurations personnalisables.

> **Note**
>
> Avec des configurations temporelles telles que les paramètres `time`, `interval` et `window`, l'unité par défaut lorsqu'aucune unité n'est spécifiée est la milliseconde. Une valeur de `30` sans suffixe est donc égale à 30 millisecondes. Pour les autres unités de temps, incluez le suffixe approprié `s` pour les secondes, `m` pour les minutes ou `h` pour les heures.

Rubriques

- Configurations de boucle de l'outil de mise à l'échelle automatique
- Agrégation des métriques et configurations de l'historique
- Configurations au niveau des vertex de tâche
- Configurations de traitement du backlog
- Configurations des opérations de mise à l'échelle

Configurations de boucle de l’outil de mise à l’échelle automatique

L’outil de mise à l’échelle automatique récupère les métriques au niveau des vertex de la tâche pour chaque intervalle de temps configurable, les convertit en variables d’échelle, estime le parallélisme des nouveaux vertex de la tâche et le recommande au planificateur de tâches. Les métriques ne sont collectées qu’à l’issue du redémarrage de la tâche et de l’intervalle de stabilisation du cluster.

Clé de configuration	Valeur par défaut	Description	Exemples de valeur
`job.autoscaler.enabled`	`false`	Activez la mise à l’échelle automatique sur votre cluster Flink.	`true`, `false`
`job.autoscaler.decision.interval`	`60s`	Intervalle de décision de l’outil de mise à l’échelle automatique.	`30` (l’unité par défaut est la milliseconde), `5m`, `1h`
`job.autoscaler.restart.time`	`3m`	Temps de redémarrage prévu à utiliser jusqu’à ce que l’opérateur puisse le déterminer de manière fiable à partir de l’historique.	`30` (l’unité par défaut est la milliseconde), `5m`, `1h`
`job.autoscaler.stabilization.interval`	`300s`	Période de stabilisation au cours de laquelle aucune nouvelle mise à l’échelle ne sera exécutée.	`30` (l’unité par défaut est la milliseconde), `5m`, `1h`
`job.autoscaler.debug.logs.interval`	`300s`	Intervalle entre les journaux de débogage de l’outil de mise à l’échelle automatique.	`30` (l’unité par défaut est la milliseconde), `5m`, `1h`

Agrégation des métriques et configurations de l'historique

L'outil de mise à l'échelle automatique récupère les métriques et les agrège sur une fenêtre temporelle coulissante. Les métriques sont alors évaluées dans le cadre de décisions de mise à l'échelle. L'historique des décisions de mise à l'échelle pour chaque vertex de tâche sert à estimer le nouveau parallélisme. Les métriques présentent une date d'expiration basée sur l'ancienneté et une taille d'historique (1 au minimum).

Clé de configuration	Valeur par défaut	Description	Exemples de valeur
`job.autoscaler.metrics.window`	`600s`	`Scaling metrics aggregation window size.`	`30` (l'unité par défaut est la milliseconde), `5m`, `1h`
`job.autoscaler.history.max.count`	`3`	Nombre maximal de décisions de mise à l'échelle passées à conserver par vertex.	`1` sur `Integer.MAX_VALUE`
`job.autoscaler.history.max.age`	`24h`	Nombre minimal de décisions de mise à l'échelle passées à conserver par vertex.	`30` (l'unité par défaut est la milliseconde), `5m`, `1h`

Configurations au niveau des vertex de tâche

Le parallélisme de chaque vertex de tâche est modifié en fonction de l'utilisation de la cible et est limité par les limites minimale et maximale de parallélisme. Nous vous déconseillons de définir une cible d'utilisation proche de 100 % (c'est-à-dire une valeur de 1). La limite d'utilisation joue le rôle de tampon pour gérer les fluctuations de charge intermédiaires.

Clé de configuration	Valeur par défaut	Description	Exemples de valeur
`job.autoscaler.target.utilization`	`0.7`	Utilisation des vertex cibles.	`0` - `1`

Clé de configuration	Valeur par défaut	Description	Exemples de valeur
`job.autoscaler.target.utilization.boundary`	`0.4`	Limite d'utilisation des vertex cibles. La mise à l'échelle ne sera pas effectuée si le taux de traitement actuel est compris entre `[target_rate / (target_utilization - boundary)` et `(target_rate / (target_utilization + boundary)]` .	`0 - 1`
`job.autoscaler.vertex.min-parallelism`	`1`	Parallélisme minimal que l'outil de mise à l'échelle automatique peut utiliser.	`0 - 200`
`job.autoscaler.vertex.max-parallelism`	`200`	Parallélisme maximal que l'outil de mise à l'échelle automatique peut utiliser. Notez que l'outil de mise à l'échelle automatique ignore cette limite si elle est supérieure au parallélisme maximal configuré dans la configuration Flink ou directement sur chaque opérateur.	`0 - 200`

Configurations de traitement du backlog

Le vertex de la tâche requiert des ressources supplémentaires pour gérer les événements en attente, ou backlogs, qui s'accumulent pendant le délai de mise à l'échelle. C'est ce que l'on appelle également la durée de `catch-up`. Si le délai de traitement du backlog dépasse la `lag -threshold` valeur configurée, l'utilisation cible du vertex de la tâche augmente jusqu'au niveau maximum. Cela permet d'éviter les opérations de mise à l'échelle inutiles pendant le traitement du backlog.

Clé de configuration	Valeur par défaut	Description	Exemples de valeur
`job.autoscaler.backlog-processing.lag-threshold`	`5m`	Seuil de retard qui évitera les mises à l'échelle inutiles tout en supprimant les messages en attente responsables du retard.	`30` (l'unité par défaut est la milliseconde), `5m`, `1h`
`job.autoscaler.catch-up.duration`	`15m`	Durée cible pour le traitement complet de tout backlog après une opération de mise à l'échelle. Définissez cette valeur sur 0 pour désactiver la mise à l'échelle basée sur le backlog.	`30` (l'unité par défaut est la milliseconde), `5m`, `1h`

Configurations des opérations de mise à l'échelle

Aucune opération de réduction n'est effectuée par l'outil de mise à l'échelle automatique immédiatement après une opération de mise à l'échelle pendant le délai de grâce. Cela permet d'éviter les cycles inutiles d' up-down-up-downopérations de balance provoqués par des fluctuations de charge temporaires.

Nous pouvons utiliser le ratio de réduction d'échelle pour réduire progressivement le parallélisme et libérer des ressources pour faire face à des pics de charge temporaires. Cela permet également d'éviter des opérations d'augmentation d'échelle mineures et inutiles après une opération de réduction majeure.

Nous pouvons détecter une opération de mise à l'échelle inefficace sur la base de l'historique des décisions d'augmentation d'échelle des vertex de tâches passées afin d'empêcher toute nouvelle modification du parallélisme.

Clé de configuration	Valeur par défaut	Description	Exemples de valeur
`job.autoscaler.scale-up.grace-period`	`1h`	Durée pendant laquelle aucune réduction d'échelle d'un vertex n'est autorisée après l'augmentation d'échelle de ce vertex.	`30` (l'unité par défaut est la milliseconde), `5m`, `1h`
`job.autoscaler.scale-down.max-factor`	`0.6`	Facteur maximal de réduction d'échelle. Une valeur égale à `1` signifie qu'aucune limite de réduction d'échelle n'est définie ; `0.6` indique que seule une réduction d'échelle de 60 % du parallélisme d'origine peut être appliquée à la tâche.	`0` - `1`
`job.autoscaler.scale-up.max-factor`	`100000.`	Ratio maximal d'augmentation d'échelle. Une valeur de `2.0` signifie que seule une augmentation d'échelle de 200 % du parallélisme actuel peut être appliquée à la tâche.	`0` - `Integer.MAX_VALUE`
`job.autoscaler.scaling.effectiveness.detection.enabled`	`false`	Permet d'activer la détection des opérations de mise à l'échelle inefficaces et d'autoriser le blocage de nouvelles mises à l'échelle par l'outil de mise à l'échelle automatique.	`true`, `false`

Optimisation des temps de redémarrage pour les opérations de récupération et de mise à l'échelle de tâches

Lorsqu'une tâche échoue ou qu'une opération de mise à l'échelle est en cours, Flink tente de réexécuter la tâche à partir du dernier point de contrôle terminé. L'exécution du processus de redémarrage peut durer une minute ou plus, en fonction de la taille de l'état du point de contrôle et du nombre de tâches parallèles. Pendant la période de redémarrage, les tâches en attente peuvent s'accumuler pour la tâche. Flink peut cependant permettre d'optimiser la vitesse de récupération et de redémarrage des graphes d'exécution afin d'améliorer la stabilité des tâches.

Cette page décrit comment Amazon EMR Flink peut améliorer le temps de redémarrage des tâches lors des opérations de récupération ou de mise à l'échelle de tâches.

Rubriques

- Récupération locale des tâches
- Point de contrôle incrémentiel générique basé sur les journaux
- Récupération précise
- Mécanisme de redémarrage combiné dans le planificateur adaptatif

Récupération locale des tâches

> **Note**
>
> La récupération locale des tâches est prise en charge avec Amazon EMR 6.0.0 et versions ultérieures.

Avec les points de contrôle Flink, chaque tâche crée un instantané de son état, que Flink écrit sur un système de stockage distribué tel qu'Amazon S3. En cas de récupération, les tâches restaurent leur état à partir du stockage distribué. Le stockage distribué offre une tolérance aux pannes et peut redistribuer l'état lors de la mise à l'échelle, car tous les nœuds peuvent y accéder.

Cependant, un magasin distribué à distance présente également un inconvénient : toutes les tâches doivent lire leur état depuis un emplacement distant sur le réseau, ce qui peut entraîner l'augmentation du temps de récupération pour les états importants lors des opérations de récupération ou de mise à l'échelle de tâches.

La récupération locale des tâches permet de résoudre ce problème. Les tâches enregistrent leur état au point de contrôle sur un stockage secondaire local à la tâche, par exemple sur un disque local. Elles stockent également leur état sur le stockage principal, à savoir Amazon S3 dans notre cas. Pendant la récupération, le planificateur planifie les tâches sur le même Task Manager que celui dans lequel les tâches ont été exécutées précédemment afin qu'elles puissent être récupérées depuis le magasin d'état local au lieu de lire depuis le magasin d'état distant. Pour plus d'informations, voir la rubrique Récupération locale des tâches de la documentation Apache Flink.

Nos tests d'évaluation avec des exemples de tâches ont montré que le temps de récupération était passé de quelques minutes à quelques secondes grâce à l'activation de la récupération locale des tâches.

Pour activer la récupération locale des tâches, définissez les configurations suivantes dans votre fichier `flink-conf.yaml`. Spécifiez la valeur de l'intervalle de point de contrôle en millisecondes.

```
state.backend.local-recovery: true
state.backend: hasmap or rocksdb
state.checkpoints.dir: s3://storage-location-bucket-path/checkpoint
execution.checkpointing.interval: 15000
```

Point de contrôle incrémentiel générique basé sur les journaux

> **Note**
>
> Le point de contrôle incrémentiel générique basé sur les journaux est pris en charge avec Amazon EMR 6.10.0 et versions ultérieures.

Le point de contrôle incrémentiel générique basé sur les journaux a été ajouté dans Flink 1.16 pour accélérer les points de contrôle. Un intervalle de point de contrôle plus court entraîne souvent une réduction du travail de récupération, car moins d'événements doivent être traités de nouveau après la récupération. Pour plus d'informations, accédez à la page Improving speed and stability of checkpointing with generic log-based incremental checkpoints sur le blog Apache Flink.

Sur base de quelques exemples de tâches, nos tests d'évaluation ont montré que le temps de contrôle était passé de quelques minutes à quelques secondes grâce au point de contrôle incrémentiel générique basé sur les journaux.

Pour activer les points de contrôle incrémentiels génériques basés sur les journaux, définissez les configurations suivantes dans votre fichier `flink-conf.yaml`. Spécifiez la valeur de l'intervalle de point de contrôle en millisecondes.

```
state.backend.changelog.enabled: true
state.backend.changelog.storage: filesystem
dstl.dfs.base-path: s3://bucket-path/changelog
state.backend.local-recovery: true
state.backend: rocksdb
state.checkpoints.dir: s3://bucket-path/checkpoint
execution.checkpointing.interval: 15000
```

Récupération précise

> **Note**
>
> La prise en charge de la récupération précise pour le planificateur par défaut est disponible avec Amazon EMR 6.0.0 et versions ultérieures. La prise en charge de la récupération précise pour le planificateur adaptatif est disponible avec Amazon EMR 6.15.0 et versions ultérieures.

Lorsqu'une tâche échoue pendant son exécution, Flink réinitialise l'intégralité du graphe d'exécution et déclenche une réexécution complète à partir du dernier point de contrôle terminé. Cette opération est plus chère qu'une simple réexécution des tâches qui ont échoué. La récupération précise redémarre uniquement le composant connecté au pipeline de la tâche ayant échoué. Dans l'exemple suivant, le graphe de tâches présente 5 sommets (`A` à `E`). Toutes les connexions entre les sommets sont en pipeline avec une distribution ponctuelle, et la valeur de `parallelism.default` pour la tâche est définie sur 2.

```
A # B # C # D # E
```

Dans cet exemple, 10 tâches sont en cours d'exécution au total. Le premier pipeline (`a1toe1`) fonctionne sur un TaskManager (`TM1`), et le second pipeline (`a2toe2`) fonctionne sur un autre TaskManager (`TM2`).

```
a1 # b1 # c1 # d1 # e1
a2 # b2 # c2 # d2 # e2
```

Deux composants sont connectés en pipeline : `a1 # e1` et `a2 # e2`. En cas d'`TM2`échec `TM1` ou d'échec, l'échec n'a d'impact que sur les 5 tâches du pipeline où TaskManager était en train de courir. La stratégie de redémarrage démarre uniquement le composant en pipeline concerné.

La récupération précise ne fonctionne qu'avec des tâches Flink parfaitement parallèles. Elle n'est pas prise en charge par `keyBy()` ou par les opérations `redistribute()`. Pour plus d'informations, accédez à la page FLIP-1 : Fine Grained Recovery from Task Failures du projet Jira Flink Improvement Proposal.

Pour activer la récupération précise, définissez les configurations suivantes dans votre fichier `flink-conf.yaml`.

```
jobmanager.execution.failover-strategy: region
restart-strategy: exponential-delay or fixed-delay
```

Mécanisme de redémarrage combiné dans le planificateur adaptatif

> **Note**
>
> Le mécanisme de redémarrage combiné du planificateur adaptatif est pris en charge par Amazon EMR 6.15.0 et versions ultérieures.

Le planificateur adaptatif peut ajuster le parallélisme de la tâche en fonction des emplacements disponibles. Si le nombre d'emplacements disponibles est insuffisant, le planificateur réduit automatiquement le parallélisme pour s'adapter au parallélisme des tâches configuré. Si de nouveaux emplacements sont disponibles, la tâche fait l'objet d'une augmentation d'échelle selon le parallélisme des tâches configuré. Un planificateur adaptatif permet d'éviter les temps d'arrêt de la tâche lorsque les ressources disponibles sont insuffisantes. Il s'agit du planificateur pris en charge par l'outil de mise à l'échelle automatique Flink. Nous recommandons donc l'utilisation d'un planificateur adaptatif avec Amazon EMR Flink. Les planificateurs adaptatifs peuvent toutefois effectuer plusieurs redémarrages en peu de temps, à raison d'un redémarrage pour chaque nouvelle ressource ajoutée, ce qui peut entraîner une baisse des performances de la tâche.

Avec Amazon EMR 6.15.0 et versions ultérieures, Flink dispose d'un mécanisme de redémarrage combiné dans le planificateur adaptatif qui ouvre une fenêtre de redémarrage lorsque la première ressource est ajoutée, puis attend l'intervalle de fenêtre configuré de 1 minute par défaut. Un seul redémarrage est effectué lorsque les ressources disponibles sont suffisantes pour exécuter la tâche avec un parallélisme configuré ou lorsque l'intervalle expire.

Grâce à quelques exemples de tâches, nos tests d'évaluation ont montré que cette fonctionnalité traite 10 % d'enregistrements supplémentaires par rapport au comportement par défaut lorsque vous utilisez le planificateur adaptatif et l'outil de mise à l'échelle automatique Flink.

Pour activer le mécanisme de redémarrage combiné, définissez les configurations suivantes dans votre fichier `flink-conf.yaml`.

```
jobmanager.adaptive-scheduler.combined-restart.enabled: true
jobmanager.adaptive-scheduler.combined-restart.window-interval: 1m
```

Travailler avec les jobs Flink de Zeppelin dans Amazon EMR

Introduction

Les versions 6.10.0 et supérieures d'Amazon EMR prennent en charge l'intégration Apache Zeppelin avec Apache Flink. Vous pouvez soumettre des jobs Flink de manière interactive via les blocs-notes Zeppelin. Avec l'interpréteur Flink, vous pouvez exécuter des requêtes Flink, définir des tâches de streaming et de traitement par lots Flink, et visualiser le résultat dans les blocs-notes Zeppelin. L'interpréteur Flink est basé sur Flink REST API. Cela vous permet d'accéder aux tâches Flink et de les manipuler depuis l'environnement Zeppelin pour effectuer un traitement et une analyse des données en temps réel.

Il existe quatre sous-interprètes dans Flink Interpreter. Ils ont des objectifs différents, mais ils se trouvent tous dans la JVM et partagent les mêmes points d'entrée préconfigurés vers Flink (`ExecutionEnviroment`, `StreamExecutionEnvironment`, `BatchTableEnvironment`, `StreamTableEnvironment`). Les interprètes sont les suivants :

- `%flink` – Crée `ExecutionEnvironment`, `StreamExecutionEnvironment`, `BatchTableEnvironment`, `StreamTableEnvironment`, et fournit un environnement Scala
- `%flink.pyflink` – Fournit un environnement Python
- `%flink.ssql` – Fournit un environnement SQL de streaming
- `%flink.bsql` – Fournit un environnement SQL par lots

Prérequis

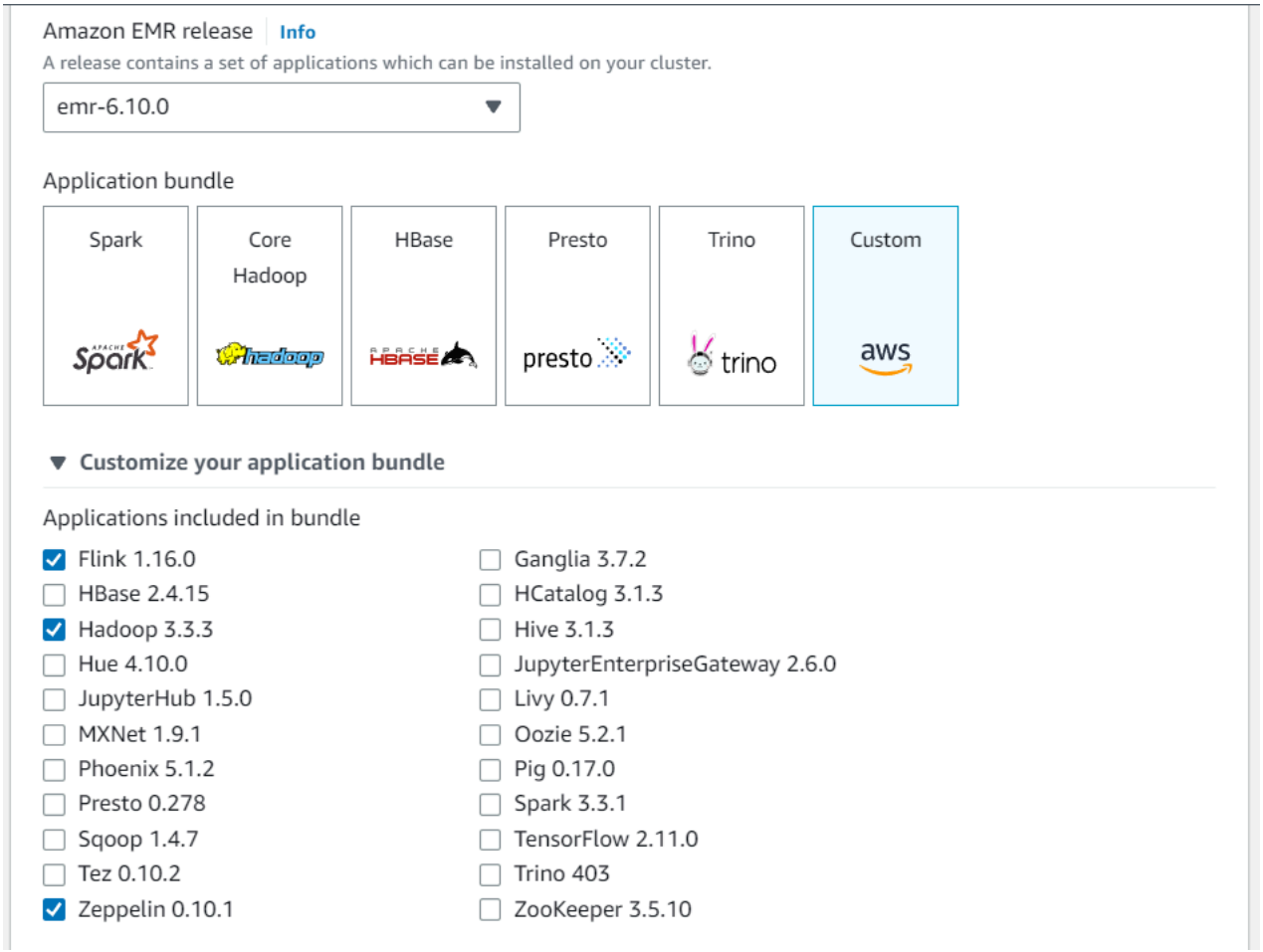
- L'intégration de Zeppelin à Flink est prise en charge pour les clusters créés avec Amazon EMR 6.10.0 et versions ultérieures.

- Pour afficher les interfaces Web hébergées sur des clusters EMR conformément à ces étapes, vous devez configurer un tunnel SSH pour autoriser l'accès entrant. Pour plus d'informations, consultez la rubrique Configuration des paramètres de proxy pour afficher les sites web hébergés sur le nœud primaire.

Configurer Zeppelin-Flink sur un cluster EMR

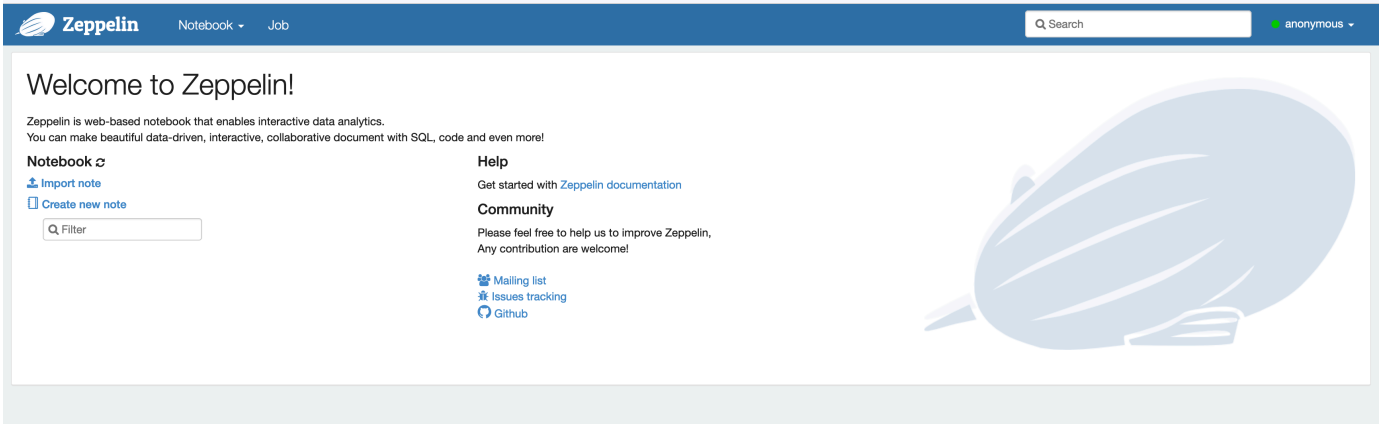
Procédez comme suit pour configurer Apache Flink sur Apache Zeppelin afin qu'il s'exécute sur un cluster EMR :

1. Créez un nouveau cluster depuis la console Amazon EMR. Sélectionnez emr-6.10.0 ou supérieur pour la version Amazon EMR. Choisissez ensuite de personnaliser votre bundle d'applications avec l'option Personnaliser. Incluez au moins Flink, Hadoop et Zeppelin dans votre offre groupée.

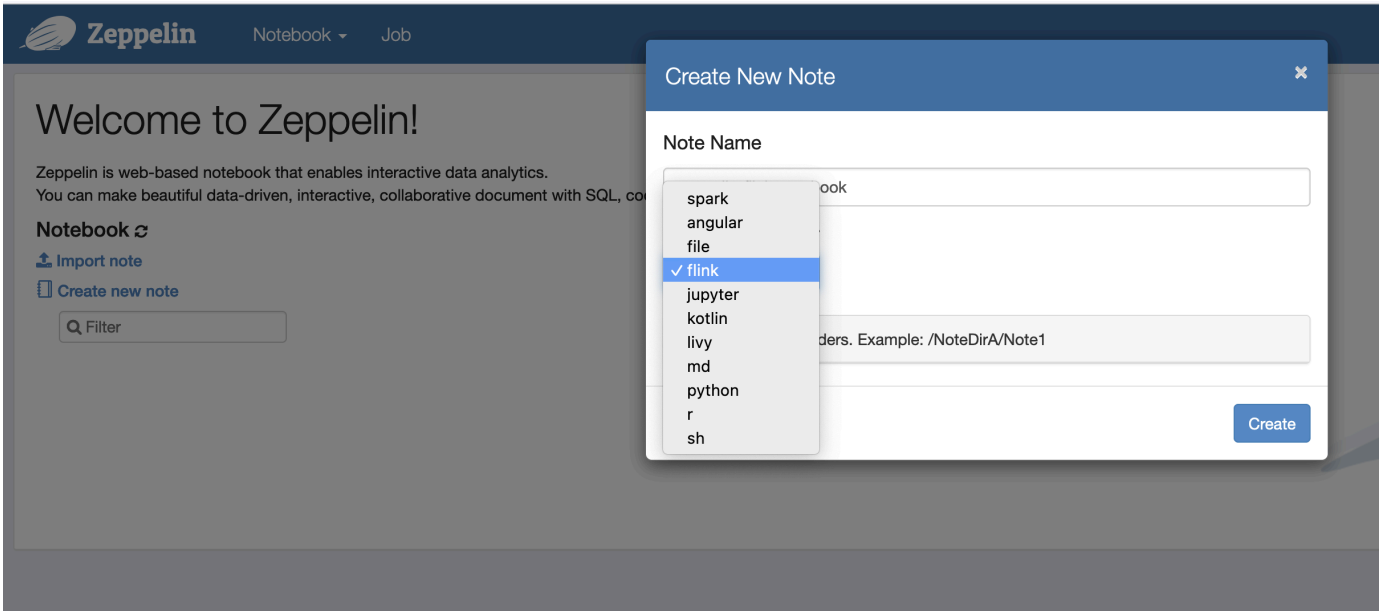


2. Créez le reste de votre cluster avec les paramètres que vous préférez.

3. Une fois que votre cluster est en cours d'exécution, sélectionnez-le dans la console pour afficher ses détails et ouvrez l'onglet Applications. Sélectionnez Zeppelin dans la section Interfaces utilisateur de l'application pour ouvrir l'interface Web Zeppelin. Assurez-vous d'avoir configuré l'accès à l'interface Web Zeppelin avec un tunnel SSH vers le nœud primaire et une connexion proxy, comme décrit dans le Prérequis.



4. Vous pouvez désormais créer une nouvelle note dans un carnet Zeppelin avec Flink comme interpréteur par défaut.



5. Reportez-vous aux exemples de code suivants qui montrent comment exécuter des tâches Flink à partir d'un bloc-notes Zeppelin.

Exécuter des tâches Flink avec Zeppelin-Flink sur un cluster EMR

- Exemple 1, Flink Scala

a) WordCount Exemple de lot (SCALA)

```
%flink

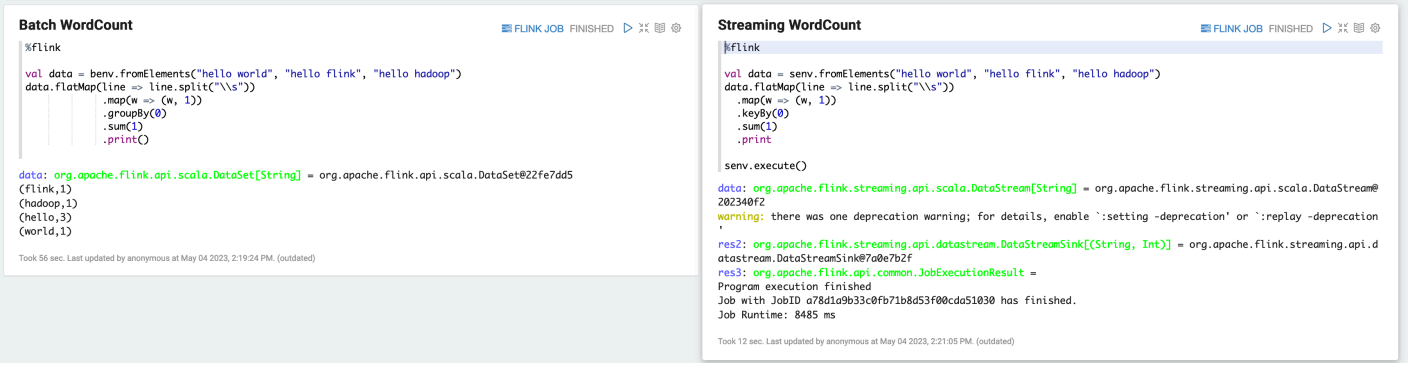
val data = benv.fromElements("hello world", "hello flink", "hello hadoop")
data.flatMap(line => line.split("\\s"))
              .map(w => (w, 1))
              .groupBy(0)
              .sum(1)
              .print()
```

b) WordCount Exemple de streaming (SCALA)

```
%flink

val data = senv.fromElements("hello world", "hello flink", "hello hadoop")
data.flatMap(line => line.split("\\s"))
  .map(w => (w, 1))
  .keyBy(0)
  .sum(1)
  .print

senv.execute()
```



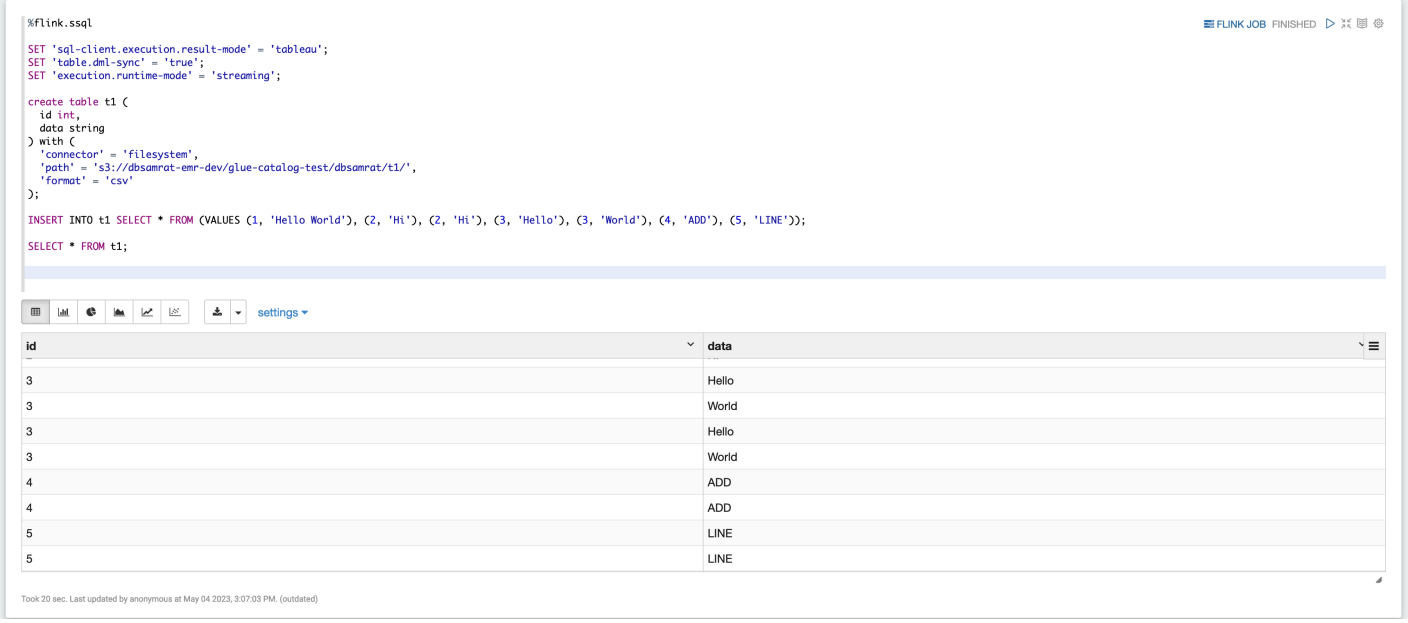
- Exemple 2, Flink Streaming SQL

```
%flink.ssql
SET 'sql-client.execution.result-mode' = 'tableau';
SET 'table.dml-sync' = 'true';
SET 'execution.runtime-mode' = 'streaming';
```

```
create table dummy_table (
  id int,
  data string
) with (
  'connector' = 'filesystem',
  'path' = 's3://s3-bucket/dummy_table',
  'format' = 'csv'
);

INSERT INTO dummy_table SELECT * FROM (VALUES (1, 'Hello World'), (2, 'Hi'), (2,
 'Hi'), (3, 'Hello'), (3, 'World'), (4, 'ADD'), (5, 'LINE'));

SELECT * FROM dummy_table;
```



- Exemple 3, Pyflink. Notez que vous devez télécharger votre propre exemple de fichier texte nommé `word.txt` dans votre compartiment S3.

```
%flink.pyflink

import argparse
import logging
import sys

from pyflink.common import Row
from pyflink.table import (EnvironmentSettings, TableEnvironment, TableDescriptor,
 Schema,
                           DataTypes, FormatDescriptor)
```

```
from pyflink.table.expressions import lit, col
from pyflink.table.udf import udtf

def word_count(input_path, output_path):
    t_env = TableEnvironment.create(EnvironmentSettings.in_streaming_mode())
    # write all the data to one file
    t_env.get_config().set("parallelism.default", "1")

    # define the source
    if input_path is not None:
        t_env.create_temporary_table(
            'source',
            TableDescriptor.for_connector('filesystem')
                           .schema(Schema.new_builder()
                                   .column('word', DataTypes.STRING())
                                   .build())
                           .option('path', input_path)
                           .format('csv')
                           .build())
        tab = t_env.from_path('source')
    else:
        print("Executing word_count example with default input data set.")
        print("Use --input to specify file input.")
        tab = t_env.from_elements(map(lambda i: (i,), word_count_data),
                                  DataTypes.ROW([DataTypes.FIELD('line',
 DataTypes.STRING())]))

    # define the sink
    if output_path is not None:
        t_env.create_temporary_table(
            'sink',
            TableDescriptor.for_connector('filesystem')
                           .schema(Schema.new_builder()
                                   .column('word', DataTypes.STRING())
                                   .column('count', DataTypes.BIGINT())
                                   .build())
                           .option('path', output_path)
                           .format(FormatDescriptor.for_format('canal-json')
                                   .build())
                           .build())
    else:
        print("Printing result to stdout. Use --output to specify output path.")
        t_env.create_temporary_table(
            'sink',
```

```
            TableDescriptor.for_connector('print')
                           .schema(Schema.new_builder()
                                   .column('word', DataTypes.STRING())
                                   .column('count', DataTypes.BIGINT())
                                   .build())
                           .build())

    @udtf(result_types=[DataTypes.STRING()])
    def split(line: Row):
        for s in line[0].split():
            yield Row(s)

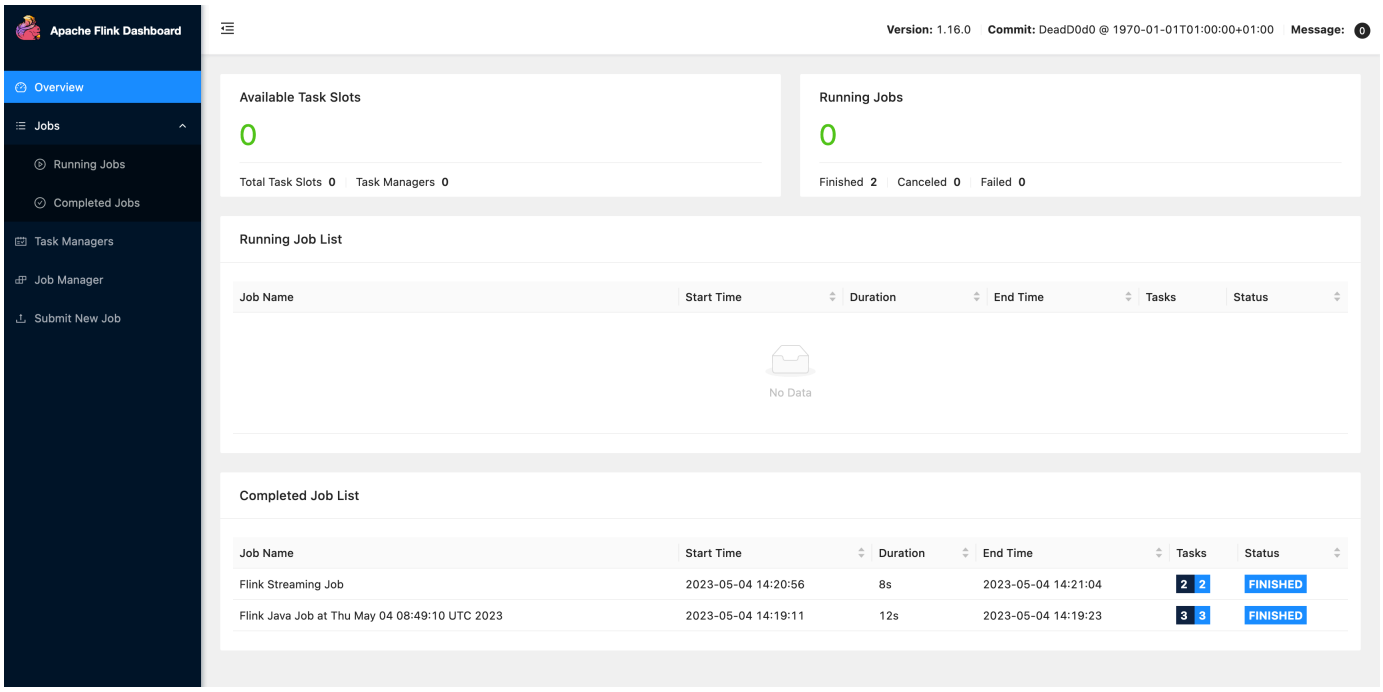
    # compute word count
    tab.flat_map(split).alias('word') \
       .group_by(col('word')) \
       .select(col('word'), lit(1).count) \
       .execute_insert('sink') \
       .wait()


logging.basicConfig(stream=sys.stdout, level=logging.INFO, format="%(message)s")


word_count("s3://s3_bucket/word.txt", "s3://s3_bucket/demo_output.txt")
```

1. Choisissez FLINK JOB dans l'interface utilisateur Zeppelin pour accéder à l'interface utilisateur Web de Flink et l'afficher.

2. Lorsque vous choisissez FLINK JOB, vous accédez à la console Web Flink dans un autre onglet de votre navigateur.



Historique des publications de Flink

Le tableau suivant répertorie la version de Flink incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version de Flink

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-7.9.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-ya

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
		rn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7,8.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7,7.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-7.6.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.5.0	1.19.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-7.4.0	1.19.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.3.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-7.2.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-5.36.2	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-7.1.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-7.0.0	1.18.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.15.0	1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-6.14.0	1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.13.0	1.17.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-6.12.0	1.17.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.11.1	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-6.11.0	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.10.1	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.10.0	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-6.9.1	1.15.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.9.0	1.15.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.8.1	1.15.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-6.8.0	1.15.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.7.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-5.36.1	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.36.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.6.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-5.35.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-6.5.0	1.14.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.4.0	1.13.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.3.1	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-6.3.0	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.2.1	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.2.0	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-6.1.1	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-6.1.0	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.34.0	1.13.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.33.1	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.33.0	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.32.1	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.32.0	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.31.1	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.31.0	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.30.2	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.30.1	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.30.0	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
mr-5.29.0	1.9.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.28.1	1.9.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.28.0	1.9.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.27.1	1.8.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.27.0	1.8.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.26.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.25.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.24.1	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.24.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.23.1	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.23.0	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.22.0	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.21.2	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.21.1	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.21.0	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.20.1	1.6.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.20.0	1.6.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.19.1	1.6.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.19.0	1.6.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.18.1	1.6.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.18.0	1.6.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.17.2	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.17.1	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.17.0	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.16.1	1.5.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.16.0	1.5.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.15.1	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.15.0	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.14.2	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.14.1	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.14.0	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.13.1	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.13.0	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.12.3	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.12.2	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.12.1	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.12.0	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.4	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.3	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.11.2	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.10.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.10.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.9.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.9.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.3	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.2	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.8.1	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.0	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.7.1	1.3.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.7.0	1.3.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.6.1	1.2.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.6.0	1.2.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.5.4	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.3	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.2	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.5.1	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.0	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.4.1	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.4.0	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.3.2	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.3.1	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.3.0	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.3	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.2	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.2.1	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.0	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.1.1	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Étiquette de version Amazon EMR	Version Flink	Composants installés avec Flink
emr-5.1.0	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Notes de mise à jour de Flink par version

Consultez les sections suivantes pour les notes de mise à jour complètes.

Amazon EMR 7.9.0 - Notes de mise à jour de Flink

Amazon EMR 7.9.0 - Modifications apportées à Flink

Type	Description
Nouvelle fonction	À partir d'Amazon EMR 7.9.0, Apache Flink prend en out-of-the-box charge les formats de fichier Avro, Parquet et ORC. Vous pouvez utiliser ces formats directement avec n'importe quelle API Flink (DataStream, Table ou SQL) sans nécessiter de configuration supplémen taire.
Nouvelle fonction	À partir de la version 7.9.0 d'Amazon EMR, vous pouvez activer plus facilement les catalogues de données Hive Metastore ou AWS Glue en utilisant les paramètres de configuration. Ajoutez l'une `hive.enabled: true` ou `glue.enabled: true` l'autre `flink-conf` classification lors de la création

Type	Description
	du cluster pour configurer automatiquement le catalogue de données correspondant. Cette approche rationalisée élimine les étapes de configuration manuelle qui étaient auparavant requises.

Amazon EMR 7.8.0 - Notes de mise à jour de Flink

Configuration — EMR Flink fonctionne immédiatement avec le S3A dans toutes les régions/partitions. AWS

Amazon EMR 7.7.0 - Notes de mise à jour de Flink

- Le shell SQL Flink peut être invoqué facilement à l'aide de la commande `flink-sql-client` liée symboliquement à `/usr/lib/flink/bin/sql-client.sh`

Amazon EMR 7.6.0 - Notes de mise à jour de Flink

Amazon EMR 7.6.0 - Fonctionnalités de Flink

- Aucune modification pour la version.

Amazon EMR 7.5.0 - Notes de mise à jour de Flink

Type	Description
Fonctionnalité	Ajout de la prise en charge de l'exécution de tâches Flink avec un fichier JAR distant.
Amélioration	Sécurisez le thread d'exclusion et d'inclusion des sommets.

Amazon EMR 7.5.0 - Fonctionnalités de Flink

- À partir d'Amazon EMR 7.5.0, vous pouvez spécifier un emplacement Amazon S3 comme chemin JAR lorsque vous utilisez les commandes `run` et Apache `run-application` Flink CLI. Lorsque vous fournissez un chemin S3, EMR télécharge automatiquement le fichier JAR depuis Amazon S3 vers le stockage EBS du cluster. Chaque fois que vous spécifiez le même fichier JAR, EMR télécharge la dernière version depuis Amazon S3 au lieu de réutiliser le fichier JAR existant sur le cluster.
- À partir d'Amazon EMR 7.5.0, les clients peuvent transmettre le chemin distant (un emplacement S3) en tant que chemin JAR avec les commandes `run` Flink CLI`run-application`. Le fichier JAR est ensuite automatiquement extrait du magasin S3 vers le stockage EBS du cluster. Si le même fichier JAR est à nouveau fourni, il télécharge le dernier depuis S3 et ne réutilise pas le fichier JAR existant sur le cluster.

Amazon EMR 7.4.0 - Notes de mise à jour de Flink

Type	Description
Upgrade	La version de Flink est mise à niveau vers la version 1.19.1.

Amazon EMR 7.3.0 - Notes de mise à jour de Flink

- Par défaut, les clusters dont le chiffrement en transit est activé via la configuration de sécurité utiliseront le protocole TLS 1.3 pour la communication interne entre les processus Flink, le point de terminaison REST de Job Manager et le serveur d'historique des tâches Flink.

Amazon EMR 7.2.0 - Notes de mise à jour de Flink

Type	Description
Amélioration	Support pour l'ajout d'étiquettes personnal isées au service Kubernetes par tâche Flink via la configuration. `kubernetes.service .labels`

Ganglia

> **Note**
>
> La dernière version d'Amazon EMR à inclure Ganglia était Amazon EMR 6.15.0. Pour surveiller votre cluster, les versions supérieures à 6.15.0 incluent l'agent Amazon CloudWatch .

Le projet open source Ganglia est un système distribué évolutif, conçu pour surveiller les clusters et les grilles tout en réduisant au maximum l'impact sur leurs performances. Lorsque vous activez Ganglia sur votre cluster, vous pouvez générer des rapports et afficher la performance du cluster dans son ensemble, ainsi qu'inspecter la performance des instances de chaque nœud. Ganglia est également configuré pour intégrer et visualiser les métriques Hadoop et Spark. Pour plus d'informations sur le projet open source Ganglia, accédez à http://ganglia.info/.

Lorsque vous affichez l'interface utilisateur web de Ganglia dans un navigateur, vous voyez une vue d'ensemble des performances du cluster, avec des graphiques détaillant la charge, l'utilisation de la mémoire, l'utilisation de l'UC et le trafic réseau du cluster. Sous les statistiques du cluster, des graphiques sont fournis pour chaque serveur individuel figurant dans le cluster.

Le tableau suivant répertorie la version de Ganglia incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Ganglia.

Pour connaître la version des composants installés avec Ganglia dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version de Ganglia pour emr-6.15.0

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-6.15.0	Ganglia 3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
		server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, webserver

Le tableau suivant répertorie la version de Ganglia incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Ganglia.

Pour la version des composants installés avec Ganglia dans cette version, voir la version 5.36.2 Versions des composants.

Informations sur la version de Ganglia pour emr-5.36.2

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.36.2	Ganglia 3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, webserver

Rubriques

- Création d'un cluster à l'aide de Ganglia
- Affichage des métriques Ganglia
- Métriques Hadoop et Spark dans Ganglia
- Historique des versions de Ganglia

Création d'un cluster à l'aide de Ganglia

> **Note**
>
> La dernière version d'Amazon EMR à inclure Ganglia était Amazon EMR 6.15.0. Pour surveiller votre cluster, les versions supérieures à 6.15.0 incluent l'agent Amazon CloudWatch .

Pour créer un cluster avec Ganglia à l'aide de la console

1. Accédez à la nouvelle console Amazon EMR et sélectionnez Changer pour l'ancienne console depuis le menu latéral. Pour plus d'informations sur ce qu'implique le passage à l'ancienne console, consultez la rubrique Utilisation de l'ancienne console.
2. Choisissez Créer un cluster.
3. Dans Software configuration (Configuration des logiciels), choisissez All Applications (Toutes les applications), Core Hadoop ou Spark.
4. Procédez à la création du cluster avec les configurations qui conviennent.

Pour ajouter Ganglia à un cluster à l'aide du AWS CLI

Dans le AWS CLI, vous pouvez ajouter Ganglia à un cluster en utilisant `create-cluster` le `--applications` paramètre. Si vous spécifiez uniquement Ganglia à l'aide du paramètre `--applications`, Ganglia est la seule application installée.

- Tapez la commande suivante pour ajouter Ganglia lorsque vous créez un cluster et remplacez-la *myKey* par le nom de votre paire de EC2 clés.

  > **Note**
  >
  > Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Spark cluster with Ganglia" --release-label
 emr-6.15.0; \
```

```
--applications Name=Spark Name=Ganglia \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge \
--instance-count 3 --use-default-roles
```

Lorsque vous spécifiez le nombre d'instances sans utiliser le paramètre `--instance-groups`, un seul nœud maître est lancé et les instances restantes sont lancées en tant que nœuds principaux. Tous les nœuds utilisent le type d'instance spécifié dans la commande.

> **Note**
>
> Si vous n'avez pas encore créé le rôle de service EMR et le profil d' EC2instance par défaut, tapez `aws emr create-default-roles` pour les créer avant de taper la `create-cluster` sous-commande.

Pour plus d'informations sur l'utilisation des commandes Amazon EMR dans le AWS CLI, consultez. https://docs.aws.amazon.com/cli/latest/reference/emr

Affichage des métriques Ganglia

> **Note**
>
> La dernière version d'Amazon EMR à inclure Ganglia était Amazon EMR 6.15.0. Pour surveiller votre cluster, les versions supérieures à 6.15.0 incluent l'agent Amazon CloudWatch .

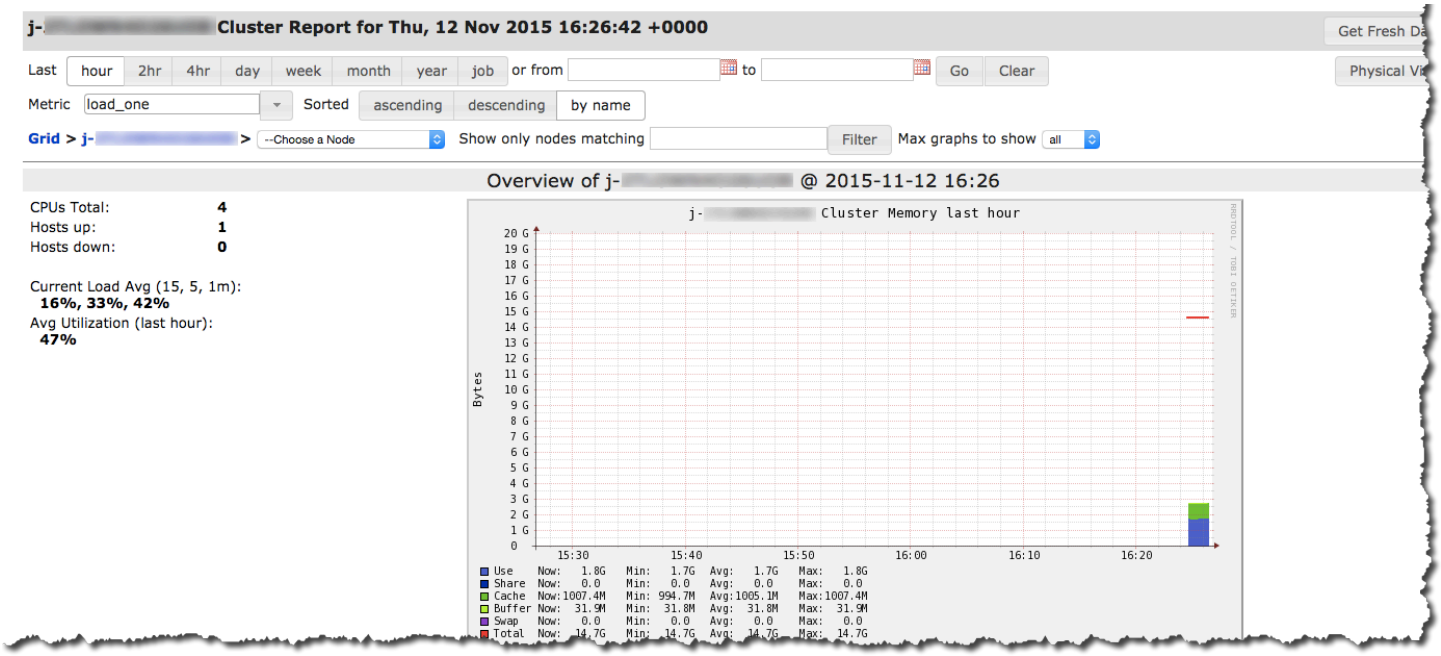
Ganglia fournit une interface utilisateur basée sur le web que vous pouvez utiliser pour afficher les métriques collectées par Ganglia. Lorsque vous exécutez Ganglia sur Amazon EMR, l'interface web s'exécute sur le nœud principal et peut être affichée à l'aide du réacheminement de port, qualifié également de création d'un tunnel SSH. Pour plus d'informations sur l'affichage des interfaces web sur Amazon EMR, consultez la section Affichage des interfaces web hébergées sur les clusters EMR dans le Guide de gestion Amazon EMR.

Pour afficher l'interface web Ganglia

1. Utilisez SSH pour accéder par tunnel au nœud maître et créer une connexion sécurisée. Pour plus d'informations sur la façon de créer un tunnel SSH pour le nœud principal,

consultez Option 2, partie 1 : Configuration d'un tunnel SSH vers le nœud principal à l'aide du réacheminement de port dynamique dans le Guide de gestion d'Amazon EMR.

2. Installez un navigateur Web avec un outil proxy, tel que le FoxyProxy plug-in pour Firefox, afin de créer un proxy SOCKS pour les domaines du type *ec2*.amazonaws.com*. Pour plus d'informations, consultez Option 2, partie 2 : Configuration des paramètres de proxy pour afficher les sites web hébergés sur le nœud principal dans le Guide de gestion d'Amazon EMR.
3. Une fois le proxy défini et la connexion SSH ouverte, vous pouvez afficher l'interface utilisateur de Ganglia en ouvrant une fenêtre de navigateur avec http ://*master-public-dns-name*/ganglia/, où se *master-public-dns-name* trouve l'adresse DNS publique du serveur principal dans le cluster EMR.



Métriques Hadoop et Spark dans Ganglia

> **Note**
>
> La dernière version d'Amazon EMR à inclure Ganglia était Amazon EMR 6.15.0. Pour surveiller votre cluster, les versions supérieures à 6.15.0 incluent l'agent Amazon CloudWatch .

Ganglia fournit des rapports sur les métriques Hadoop pour chaque instance. Les différents types de métriques sont préfixés par catégorie : système de fichiers distribué (dfs.*), machine virtuelle Java (jvm.*), (mapred.*) et appels de procédure MapReduce distante (rpc.*).

Les métriques Ganglia basées sur YARN, telles que Spark et Hadoop, ne sont pas disponibles pour les versions EMR 4.4.0 et 4.5.0. Utilisez une version ultérieure pour employer ces métriques.

Les métriques Ganglia pour Spark ont généralement des préfixes pour l'ID d'application YARN et Spark. DAGScheduler Ainsi, les préfixes suivent ce format :

- DAGScheduler.*
- application_xxxxxxxxxx_xxxx.driver.*
- application_xxxxxxxxxx_xxxx.executor.*

Historique des versions de Ganglia

> **Note**
>
> La dernière version d'Amazon EMR à inclure Ganglia était Amazon EMR 6.15.0. Pour surveiller votre cluster, les versions supérieures à 6.15.0 incluent l'agent Amazon CloudWatch .

Le tableau suivant répertorie la version de Ganglia incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations de version de Ganglia

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.36.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
		hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.15.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.14.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-6.13.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.12.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.11.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-6.11.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.10.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.10.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-6.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.9.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-6.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.36.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.36.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.35.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-6.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.3.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-6.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.2.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.2.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-6.1.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.1.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.0.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-6.0.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.34.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.33.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.33.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.32.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.32.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.31.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.31.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.30.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.30.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.30.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
mr-5.29.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.28.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.28.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.27.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.27.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.26.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.25.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.24.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.24.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.23.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.23.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.22.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.21.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.21.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.21.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.20.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.20.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.19.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.19.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.18.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.18.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.17.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.17.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.17.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.16.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.16.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.15.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.15.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.14.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.14.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.14.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.13.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.13.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.12.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.11.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.11.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.10.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.10.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.9.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.8.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.8.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.7.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.6.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.5.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.5.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.5.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.5.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.4.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.3.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.3.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.2.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.2.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.2.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.2.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.1.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.1.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.0.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.0.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.0.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-5.0.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.6	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.5	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-4.9.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-4.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.5	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-4.8.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-4.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-4.7.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-4.6.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Étiquette de version Amazon EMR	Version Ganglia	Composants installés avec Ganglia
emr-4.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.2.0	3.6.0	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Apache Hadoop

Apache Hadoop est une infrastructure logicielle Java open source qui prend en charge le traitement d'importants volumes de données sur un cluster d'instances. Il peut s'exécuter sur une seule instance ou sur des milliers d'instances. Hadoop utilise de nombreux modèles de traitement, tels que MapReduce Tez, pour répartir le traitement sur plusieurs instances et utilise également un système de fichiers distribué appelé HDFS pour stocker les données sur plusieurs instances. Hadoop surveille l'intégrité des instances dans le cluster et peut se remettre d'une défaillance d'un ou de plusieurs nœuds. De cette manière, Hadoop fournit une capacité accrue de stockage et de traitement, ainsi qu'une disponibilité élevée. Pour plus d'informations, consultez la documentation Hadoop.

Le tableau suivant répertorie la version de Hadoop incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants qu'Amazon EMR installe avec Hadoop.

Pour la version des composants installés avec Hadoop dans cette version, consultez la version 7.9.0 Versions des composants.

Informations sur la version de Hadoop pour emr-7.9.0

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-7.9.0	Hadoop 3.4.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Le tableau suivant répertorie la version de Hadoop incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Hadoop.

Pour connaître la version des composants installés avec Hadoop dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version de Hadoop pour emr-6.15.0

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.15.0	Hadoop 3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Le tableau suivant répertorie la version de Hadoop incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Hadoop.

Pour la version des composants installés avec Hadoop dans cette version, consultez la version 5.36.2 Versions des composants.

Informations sur la version de Hadoop pour emr-5.36.2

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.36.2	Hadoop 2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
		nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server

À partir d'Amazon EMR 5.18.0, vous pouvez utiliser le référentiel d'artefacts d'Amazon EMR pour générer le code de votre tâche en fonction des versions exactes des bibliothèques et des dépendances qui sont disponibles avec des versions spécifiques d'Amazon EMR. Pour de plus amples informations, veuillez consulter Vérification des dépendances à l'aide du référentiel d'artefacts d'Amazon EMR.

Rubriques

- Configuration de Hadoop
- Chiffrement transparent dans HDFS sur Amazon EMR
- Création ou exécution d'une application Hadoop
- Lire les objets restaurés
- Activer la détection non uniforme des accès à la mémoire pour les conteneurs YARN
- Emballage en poubelle YARN
- Historique des versions de Hadoop

Configuration de Hadoop

Les sections suivantes fournissent les paramètres de configuration par défaut pour HDFS, les tâches et les démons Hadoop.

Rubriques

- Configuration de la tâche
- Paramètres de configuration de démon Hadoop
- Configuration HDFS

Configuration de la tâche

Vous pouvez définir des variables de configuration pour optimiser les performances de vos MapReduce tâches. Cette section fournit les valeurs par défaut des paramètres importants. Les valeurs par défaut varient en fonction du type d' EC2 instance du nœud utilisé dans le cluster. HBase est disponible lors de l'utilisation de la version 4.6.0 ou ultérieure d'Amazon EMR. Différents paramètres par défaut sont utilisés lors de HBase l'installation. Ces valeurs sont fournies avec les valeurs par défaut initiales.

Hadoop 2 utilise deux paramètres : `mapreduce.map.java.opts` et, pour configurer la mémoire`mapreduce.reduce.java.opts`, respectivement pour cartographier et réduireJVMs . Ces paramètres remplacent l'option de configuration unique `mapreduce.map.java.opts` des versions antérieures de Hadoop.

De la même façon, `mapred.job.jvm.num.tasks` remplace `mapred.job.reuse.jvm.num.tasks` dans Hadoop 2.7.2 et versions ultérieures. Amazon EMR définit cette valeur sur 20 quel que soit le type d' EC2 instance. Vous pouvez remplacer ce paramètre à l'aide de la classification de configuration `mapred-site`. La définition d'une valeur de `-1` indique qu'une machine virtuelle Java peut être réutilisée à l'infini dans une même tâche, et une valeur de 1 indique qu'une nouvelle machine virtuelle Java est générée pour chaque tâche.

Par exemple, pour définir la valeur de `mapred.job.jvm.num.tasks` sur -1, vous pouvez créer un fichier contenant les éléments suivants :

```
[
 {
  "Classification": "mapred-site",
  "Properties": {
   "mapred.job.jvm.num.tasks": "-1"
  }
 }
]
```

Lorsque vous utilisez la `create-cluster` commande ou la `modify-instance-groups` commande du AWS CLI, vous pouvez alors référencer le fichier de configuration JSON. Dans l'exemple suivant, le fichier de configuration est enregistré sous la forme `myConfig.json` et stocké dans Amazon S3.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge \
--instance-count 3 --applications Name=Hadoop --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json \
--use-default-roles
```

De la même façon, vous pouvez modifier les valeurs par défaut répertoriées ci-dessous à l'aide de la classification de configuration `mapred-site`, et définir plusieurs valeurs et classifications de configuration à l'aide d'un seul fichier JSON. Pour de plus amples informations, veuillez consulter Configuration des applications.

Avec la version 5.21.0 et ultérieures d'Amazon EMR, vous permet de remplacer les configurations de cluster et de spécifier des classifications de configuration supplémentaires pour chaque groupe d'instances dans un cluster en cours d'exécution. Pour ce faire, utilisez la console Amazon EMR, le AWS Command Line Interface (AWS CLI) ou le AWS SDK. Pour plus d'informations, consultez Fourniture d'une configuration pour un groupe d'instances dans un cluster en cours d'exécution.

Valeurs par défaut des paramètres de configuration des tâches

> **Note**
>
> Les valeurs spécifiées pour chaque instance sont conformes à la dernière version de l'EMR. Les versions précédentes peuvent avoir des paramètres par défaut différents.

Types d'instances

- Instances c1
- Instances c3
- Instances c4
- Instances c5

- Instances c5a
- Instances c5ad
- Instances c5d
- Instances c5n
- Instances c6a
- instances c6g
- instances c6gd
- Instances c6gn
- Instances c6i
- Instances c6id
- Instances c6in
- instances C7a
- instances c7g
- Instances c7gd
- Instances c7gn
- Instances c7i
- instances c7i-flex
- instances c8g
- instances c8gd
- Instances d2
- Instances d3
- instances d3en
- instances f2
- Instances g3
- Instances g3s
- Instances g4dn
- Instances g5
- instances g6

- instances g6e
- instances gr6
- Instances h1
- Instances i2
- Instances i3
- Instances i3en
- Instances i4g
- Instances i4i
- instances i7i
- instances i7ie
- instances i8g
- Instances im4gn
- Instances is4gen
- Instances m1
- Instances m2
- Instances m3
- Instances m4
- Instances m5
- Instances m5a
- instances m5ad
- Instances m5d
- instances m5dn
- instances m5n
- instances m5zn
- Instances m6a
- Instances m6g
- instances m6gd
- Instances m6i

- Instances m6id
- instances m6idn
- instances m6in
- Instances m7a
- Instances m7g
- Instances m7gd
- Instances m7i
- Instances m7i-flex
- instances M8G
- instances m8gd
- Instances p2
- Instances p3
- instances p4d
- Instances p5
- Instances r3
- Instances r4
- Instances r5
- Instances r5a
- instances r5ad
- instances r5b
- Instances r5d
- instances r5dn
- Instances r5n
- Instances r6a
- instances r6g
- Instances r6gd
- Instances r6i
- Instances r6id
- instances r6idn

- Instances r6in
- Instances r7a
- instances r7g
- Instances r7gd
- Instances R7i
- Instances r7iz
- instances R8G
- instances r8gd
- Instances x1
- Instances x1e
- Instances x2gd
- instances x2idn
- Instances x2iedn
- instances x8g
- instances z1d

Instances c1

c1.medium

Option de configuration	Valeur par défaut	Avec HBase installé
`mapreduce.map.java.opts`	-Xmx288m	-Xmx288m
`mapreduce.java.opts`	-Xmx288m	-Xmx288m
`mapreduce.map.memory.mb`	512	512
`mapreduce.reduce.memory.mb`	512	512
`yarn.app.mapreduce.am.resource.mb`	512	512
`yarn.scheduler.minimum-allocation-mb`	1	1

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.maximum-allocation-mb	512	512
yarn.nodemanager.resource.memory-mb	1 024	512

c1.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx864m	-Xmx864m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	1 024	1 024
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2048	2560
yarn.nodemanager.resource.memory-mb	5120	2560

Instances c3

c3.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c3.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2880	2880
yarn.app.mapreduce.am.resource.mb	2880	2880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.nodemanager.resource.memory-mb	11520	5760

c3.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2880	2880
yarn.app.mapreduce.am.resource.mb	2880	2880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

c3.8.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1331m	-Xmx1331m
mapreduce.java.opts	-Xmx2662m	-Xmx2662m
mapreduce.map.memory.mb	1664	1664

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.reduce.memory.mb	3328	3328
yarn.app.mapreduce.am.resource.mb	3328	3328
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

Instances c4

c4.large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx717m	-Xmx717m
mapreduce.java.opts	-Xmx1434m	-Xmx1434m
mapreduce.map.memory.mb	896	896
mapreduce.reduce.memory.mb	1792	1792
yarn.app.mapreduce.am.resource.mb	1792	1792
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1792	896

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.nodemanager.resource.memory-mb	1792	896

c4.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c4.2.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.reduce.memory.mb	2880	2880
yarn.app.mapreduce.am.resource.mb	2880	2880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

c4.4.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2880	2880
yarn.app.mapreduce.am.resource.mb	2880	2880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

c4.8.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1183m	-Xmx1183m
mapreduce.java.opts	-Xmx2366m	-Xmx2366m
mapreduce.map.memory.mb	1479	1479
mapreduce.reduce.memory.mb	2958	2958
yarn.app.mapreduce.am.resource.mb	2958	2958
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

Instances c5

c5.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

c5.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

c5.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

c5.9xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1456m	-Xmx1456m
mapreduce.java.opts	-Xmx2912m	-Xmx2912m
mapreduce.map.memory.mb	1820	1820
mapreduce.reduce.memory.mb	3640	3640
yarn.app.mapreduce.am.resource.mb	3640	3640

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	65536	32768
yarn.nodemanager.resource.memory-mb	65536	32768

c5.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1502m	-Xmx1502m
mapreduce.java.opts	-Xmx3004m	-Xmx3004m
mapreduce.map.memory.mb	1877	1877
mapreduce.reduce.memory.mb	3754	3754
yarn.app.mapreduce.am.resource.mb	3754	3754
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30048
yarn.nodemanager.resource.memory-mb	90112	30048

c5.18xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1547m	-Xmx1547m
mapreduce.java.opts	-Xmx3094m	-Xmx3094m
mapreduce.map.memory.mb	1934	1934
mapreduce.reduce.memory.mb	3868	3868
yarn.app.mapreduce.am.resource.mb	3868	3868
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	139264	30960
yarn.nodemanager.resource.memory-mb	139264	30960

c5.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1570m	-Xmx1570m
mapreduce.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.map.memory.mb	1963	1963
mapreduce.reduce.memory.mb	3926	3926
yarn.app.mapreduce.am.resource.mb	3926	3926

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31376
yarn.nodemanager.resource.memory-mb	188416	31376

Instances c5a

c5a.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c5a.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c5a.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c5a.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c5a.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1502m	-Xmx1502m
mapreduce.java.opts	-Xmx3004m	-Xmx3004m
mapreduce.map.memory.mb	1877	1877
mapreduce.reduce.memory.mb	3754	3754
yarn.app.mapreduce.am.resource.mb	3754	3754
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30048
yarn.nodemanager.resource.memory-mb	90112	30048

c5a.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c5a.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

Instances c5ad

c5ad.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c5ad.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c5ad.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c5ad.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c5ad.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c5ad.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c5ad.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

Instances c5d

c5d.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

c5d.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

c5d.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

c5d.9xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1456m	-Xmx1456m
mapreduce.java.opts	-Xmx2912m	-Xmx2912m
mapreduce.map.memory.mb	1820	1820
mapreduce.reduce.memory.mb	3640	3640
yarn.app.mapreduce.am.resource.mb	3640	3640

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	65536	32768
yarn.nodemanager.resource.memory-mb	65536	32768

c5d.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1502m	-Xmx1502m
mapreduce.java.opts	-Xmx3004m	-Xmx3004m
mapreduce.map.memory.mb	1877	1877
mapreduce.reduce.memory.mb	3754	3754
yarn.app.mapreduce.am.resource.mb	3754	3754
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30048
yarn.nodemanager.resource.memory-mb	90112	30048

c5d.18xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1547m	-Xmx1547m
mapreduce.java.opts	-Xmx3094m	-Xmx3094m
mapreduce.map.memory.mb	1934	1934
mapreduce.reduce.memory.mb	3868	3868
yarn.app.mapreduce.am.resource.mb	3868	3868
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	139264	30960
yarn.nodemanager.resource.memory-mb	139264	30960

c5d.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1570m	-Xmx1570m
mapreduce.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.map.memory.mb	1963	1963
mapreduce.reduce.memory.mb	3926	3926
yarn.app.mapreduce.am.resource.mb	3926	3926

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31376
yarn.nodemanager.resource.memory-mb	188416	31376

Instances c5n

c5n.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1613m	-Xmx1613m
mapreduce.java.opts	-Xmx3226m	-Xmx3226m
mapreduce.map.memory.mb	2016	2016
mapreduce.reduce.memory.mb	4032	4032
yarn.app.mapreduce.am.resource.mb	4032	4032
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8064	4032
yarn.nodemanager.resource.memory-mb	8064	4032

c5n.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1613m	-Xmx1613m
mapreduce.java.opts	-Xmx3226m	-Xmx3226m
mapreduce.map.memory.mb	2016	2016
mapreduce.reduce.memory.mb	4032	4032
yarn.app.mapreduce.am.resource.mb	4032	4032
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	16128	8064
yarn.nodemanager.resource.memory-mb	16128	8064

c5n.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1741m	-Xmx1741m
mapreduce.java.opts	-Xmx3482m	-Xmx3482m
mapreduce.map.memory.mb	2176	2176
mapreduce.reduce.memory.mb	4352	4352
yarn.app.mapreduce.am.resource.mb	4352	4352

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	34816	17408
yarn.nodemanager.resource.memory-mb	34816	17408

c5n.9xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2002m	-Xmx2002m
mapreduce.java.opts	-Xmx4004m	-Xmx4004m
mapreduce.map.memory.mb	2503	2503
mapreduce.reduce.memory.mb	5006	5006
yarn.app.mapreduce.am.resource.mb	5006	5006
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

c5n.18xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2094m	-Xmx2094m
mapreduce.java.opts	-Xmx4188m	-Xmx4188m
mapreduce.map.memory.mb	2617	2617
mapreduce.reduce.memory.mb	5234	5234
yarn.app.mapreduce.am.resource.mb	5234	5234
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

Instances c6a

c6a.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6a.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6a.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6a.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6a.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6a.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6a.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6a.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

c6a.48xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

instances c6g

c6g.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6g.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6g.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6g.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6g.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6g.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

instances c6gd

c6gd.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6gd.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6gd.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6gd.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6gd.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6gd.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instances c6gn

c6gn.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6gn.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6gn.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6gn.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6gn.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6gn.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instances c6i

c6i.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6i.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6i.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6i.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6i.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6i.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6i.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6i.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instances c6id

c6id.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6id.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6id.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6id.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6id.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6id.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6id.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6id.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instances c6in

c6in.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6in.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6in.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6in.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6in.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6in.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6in.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6in.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

instances C7a

c7a.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7a.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7a.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7a.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7a.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7a.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c7a.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c7a.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

c7a.48xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

instances c7g

c7g.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7g.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7g.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7g.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7g.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7g.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instances c7gd

c7gd.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7gd.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7gd.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7gd.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7gd.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7gd.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instances c7gn

c7gn.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7gn.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7gn.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7gn.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7gn.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7gn.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instances c7i

c7i.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7i.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7i.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7i.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7i.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7i.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c7i.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c7i.48xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

instances c7i-flex

c7i-flex.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7i-flex. 2 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7i-flex 4 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7i-flex 8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7i-flex. 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7i-flex. 16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

instances c8g

c8g.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

8 cm x 2 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

8 cm x 4 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

8 cm x 8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

8 g, 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

8 g x 16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

8 g, 24 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

8 g x 48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

instances c8gd

c8gd.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1 408	1 408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c8gd.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c8gd.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

8 cm x 8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

8 cm x 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

8 cm x 16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c8g. 24 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

8 cm x 48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

Instances d2

d2.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

d2.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

d2.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

d2.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2417m	-Xmx2417m
mapreduce.java.opts	-Xmx4834m	-Xmx4834m
mapreduce.map.memory.mb	3021	3021
mapreduce.reduce.memory.mb	6042	6042
yarn.app.mapreduce.am.resource.mb	6042	6042

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30194
yarn.nodemanager.resource.memory-mb	241664	30194

Instances d3

d3.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

d3.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

d3.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

d3.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

instances d3en

d3en.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

d3en.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

d3en.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

d3en.6xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.java.opts	-Xmx5700m	-Xmx5700m
mapreduce.map.memory.mb	3563	3563
mapreduce.reduce.memory.mb	7126	7126
yarn.app.mapreduce.am.resource.mb	7126	7126
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	28496
yarn.nodemanager.resource.memory-mb	85504	28496

d3en.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

d3en.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

instances f2

F2,6 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-XMx8055m	-XMx8055m
mapreduce.java.opts	-XMx16110 m	-XMx16110 m
mapreduce.map.memory.mb	10069	10069
mapreduce.reduce.memory.mb	20138	20138
yarn.app.mapreduce.am.resource.mb	20138	20138
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	20146
yarn.nodemanager.resource.memory-mb	241664	20146

f/2,12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-XM x 8192 m	-XM x 8192 m
mapreduce.java.opts	-XMx16384m	-XMx16384m
mapreduce.map.memory.mb	10240	10240
mapreduce.reduce.memory.mb	20480	20480
yarn.app.mapreduce.am.resource.mb	20480	20480

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	20480
yarn.nodemanager.resource.memory-mb	491520	20480

f 2,48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx8294m	-Xmx8294m
mapreduce.java.opts	-Xmx16588m	-Xmx16588m
mapreduce.map.memory.mb	10368	10368
mapreduce.reduce.memory.mb	20736	20736
yarn.app.mapreduce.am.resource.mb	20736	20736
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990656	20736
yarn.nodemanager.resource.memory-mb	1990656	20736

Instances g3

g3.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

g3.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

g3.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instances g3s

g3s.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

Instances g4dn

g4dn.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

g4dn.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

g4dn.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

g4dn.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

g4dn.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

g4dn.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3 968	3 968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253 952	31744
yarn.nodemanager.resource.memory-mb	253 952	31744

Instances g5

g5.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

g5.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

g5.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

g5.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

g5.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

g5.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

g5.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

g5.48xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

instances g6

g6.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.java.opts	-Xmx4608m	-Xmx4608m
mapreduce.map.memory.mb	2880	2880
mapreduce.reduce.memory.mb	5760	5760
yarn.app.mapreduce.am.resource.mb	5760	5760
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

g 6,2 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.java.opts	-Xmx4608m	-Xmx4608m
mapreduce.map.memory.mb	2880	2880
mapreduce.reduce.memory.mb	5760	5760
yarn.app.mapreduce.am.resource.mb	5760	5760

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

g 6,4 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2662m	-Xmx2662m
mapreduce.java.opts	-XMX 5324m	-XMX 5324m
mapreduce.map.memory.mb	3328	3328
mapreduce.reduce.memory.mb	6656	6656
yarn.app.mapreduce.am.resource.mb	6656	6656
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

g 6,8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-XM 2893 m	-XM 2893 m
mapreduce.java.opts	-XM 5786 m	-XM 5786 m
mapreduce.map.memory.mb	3616	3616
mapreduce.reduce.memory.mb	7232	7232
yarn.app.mapreduce.am.resource.mb	7232	7232
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	115712	28928
yarn.nodemanager.resource.memory-mb	115712	28928

g 6,12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-XM x 2953 m	-XM x 2953 m
mapreduce.java.opts	-XM x 5906 m	-XM x 5906 m
mapreduce.map.memory.mb	3691	3691
mapreduce.reduce.memory.mb	7382	7382
yarn.app.mapreduce.am.resource.mb	7382	7382

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	177152	29512
yarn.nodemanager.resource.memory-mb	177152	29512

g 6,16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-XM x 2995 m	-XM x 2995 m
mapreduce.java.opts	-XM 5990 m	-XM 5990 m
mapreduce.map.memory.mb	3744	3744
mapreduce.reduce.memory.mb	7488	7488
yarn.app.mapreduce.am.resource.mb	7488	7488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	239616	29952
yarn.nodemanager.resource.memory-mb	239616	29952

g 6,24 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-XM x 3038 m	-XM x 3038 m
mapreduce.java.opts	-XM 6 076 m	-XM 6 076 m
mapreduce.map.memory.mb	3797	3797
mapreduce.reduce.memory.mb	7594	7594
yarn.app.mapreduce.am.resource.mb	7594	7594
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	364544	30408
yarn.nodemanager.resource.memory-mb	364544	30408

g 6,48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	737280	30720
yarn.nodemanager.resource.memory-mb	737280	30720

instances g6e

g6e.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

g6e.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

g6e.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

6 x 8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

G6E, 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

G6E, 16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

G6E, 24 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

g 6 e 48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instances gr6

gr6,4 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-XM 5786 m	-XM 5786 m
mapreduce.java.opts	-XM 11572 m	-XM 11572 m
mapreduce.map.memory.mb	7232	7232
mapreduce.reduce.memory.mb	14464	14464
yarn.app.mapreduce.am.resource.mb	14464	14464

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	115712	28928
yarn.nodemanager.resource.memory-mb	115712	28928

gr6,8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-XM 5990 m	-XM 5990 m
mapreduce.java.opts	- XM x 1 1980 m	- XM x 1 1980 m
mapreduce.map.memory.mb	7488	7488
mapreduce.reduce.memory.mb	14976	14976
yarn.app.mapreduce.am.resource.mb	14976	14976
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	239616	29952
yarn.nodemanager.resource.memory-mb	239616	29952

Instances h1

h1.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

h1.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

h1.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

h1.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3 968	3 968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253 952	31744
yarn.nodemanager.resource.memory-mb	253 952	31744

Instances i2

i2.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i2.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i2.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i2.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instances i3

i3.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i3.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i3.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i3.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i3.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instances i3en

i3en.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

i3en.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

i3en.3xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6007m	-Xmx6007m
mapreduce.java.opts	-Xmx12014m	-Xmx12014m
mapreduce.map.memory.mb	7509	7509
mapreduce.reduce.memory.mb	15018	15018
yarn.app.mapreduce.am.resource.mb	15018	15018
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

i3en.6xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6281m	-Xmx6281m
mapreduce.java.opts	-Xmx12562m	-Xmx12562m
mapreduce.map.memory.mb	7851	7851
mapreduce.reduce.memory.mb	15702	15702
yarn.app.mapreduce.am.resource.mb	15702	15702

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

i3en.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

i3en.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

Instances i4g

i4g.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i4g.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i4g.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i4g.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i4g.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instances i4i

i4i.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i4i.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i4i.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i4i.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i4i.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i4i.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

i4i.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i4i.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

instances i7i

i7i.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i7i.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i7i.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i7 x 8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i7i 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i7i 16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

i7i 24 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i7i 48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instances i7ie

i7ie.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i7ie. 2 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i7ie. 3 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5700m	-Xmx5700m
mapreduce.java.opts	-XM 1 400 m	-XM 1 400 m
mapreduce.map.memory.mb	7125	7125
mapreduce.reduce.memory.mb	14250	14250
yarn.app.mapreduce.am.resource.mb	14250	14250
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	28504
yarn.nodemanager.resource.memory-mb	85504	28504

i7ie. 6 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-XMX 5974 m	-XMX 5974 m
mapreduce.java.opts	-XM 119,48 m	-XM 119,48 m
mapreduce.map.memory.mb	7467	7467
mapreduce.reduce.memory.mb	14934	14934
yarn.app.mapreduce.am.resource.mb	14934	14934

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29860
yarn.nodemanager.resource.memory-mb	179200	29860

i7ie. 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i7ie. 18 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-XMx6155 m	-XMx6155 m
mapreduce.java.opts	-XM 12310 m	-XM 12310 m
mapreduce.map.memory.mb	7694	7694
mapreduce.reduce.memory.mb	15388	15388
yarn.app.mapreduce.am.resource.mb	15388	15388
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	553984	30792
yarn.nodemanager.resource.memory-mb	553984	30792

i7ie. 24 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i7ie. 48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instances i8g

i8g.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

8 g, 2 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

8 g x 4 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

8 g x 8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

8 g, 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

8 g, 16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

8 g, 24 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

8 g, 48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instances im4gn

im4gn.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

im4gn.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

im4gn.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

im4gn.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

im4gn.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instances is4gen

is4gen.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3514m	-Xmx3514m
mapreduce.java.opts	-Xmx7028m	-Xmx7028m
mapreduce.map.memory.mb	4393	4393
mapreduce.reduce.memory.mb	8786	8786
yarn.app.mapreduce.am.resource.mb	8786	8786

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	17572,12	8786,06
yarn.nodemanager.resource.memory-mb	17572,12	8786,06

is4gen.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3866m	-Xmx3866m
mapreduce.java.opts	-Xmx7732m	-Xmx7732m
mapreduce.map.memory.mb	4832	4832
mapreduce.reduce.memory.mb	9664	9664
yarn.app.mapreduce.am.resource.mb	9664	9664
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	38656	19328
yarn.nodemanager.resource.memory-mb	38656	19328

is4gen.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4275m	-Xmx4275m
mapreduce.java.opts	-Xmx8550m	-Xmx8550m
mapreduce.map.memory.mb	5344	5344
mapreduce.reduce.memory.mb	10688	10688
yarn.app.mapreduce.am.resource.mb	10688	10688
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32064
yarn.nodemanager.resource.memory-mb	85504	32064

is4gen.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4480m	-Xmx4480m
mapreduce.java.opts	-Xmx8960m	-Xmx8960m
mapreduce.map.memory.mb	5600	5600
mapreduce.reduce.memory.mb	11200	11200
yarn.app.mapreduce.am.resource.mb	11200	11200

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	22400
yarn.nodemanager.resource.memory-mb	179200	22400

Instances m1

m1.small

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx288m	-Xmx288m
mapreduce.java.opts	-Xmx288m	-Xmx288m
mapreduce.map.memory.mb	512	512
mapreduce.reduce.memory.mb	512	512
yarn.app.mapreduce.am.resource.mb	512	512
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	512	512
yarn.nodemanager.resource.memory-mb	1 024	512

m1.medium

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx768m	-Xmx768m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	1 024	1 024
yarn.app.mapreduce.am.resource.mb	1 024	1 024
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2048	1 024
yarn.nodemanager.resource.memory-mb	2048	1 024

m1.large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx1024m	-Xmx1024m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	1536	1536
yarn.app.mapreduce.am.resource.mb	1536	1536

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	3072	2560
yarn.nodemanager.resource.memory-mb	5120	2560

m1.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8192	6144
yarn.nodemanager.resource.memory-mb	12288	6144

Instances m2

m2.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx864m	-Xmx864m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	1 024	1 024
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	7168	7168
yarn.nodemanager.resource.memory-mb	14336	7168

m2.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1280m	-Xmx1280m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	2560	2560
yarn.app.mapreduce.am.resource.mb	2560	2560

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8192	15360
yarn.nodemanager.resource.memory-mb	30720	15360

m2.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1280m	-Xmx1280m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	2560	2560
yarn.app.mapreduce.am.resource.mb	2560	2560
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8192	30720
yarn.nodemanager.resource.memory-mb	61440	30720

Instances m3

m3.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2880	2880
yarn.app.mapreduce.am.resource.mb	2880	2880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

m3.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2880	2880
yarn.app.mapreduce.am.resource.mb	2880	2880

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

Instances m4

m4.large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

m4.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m4.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m4.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1434m	-Xmx1434m
mapreduce.java.opts	-Xmx2868m	-Xmx2868m
mapreduce.map.memory.mb	1792	1792
mapreduce.reduce.memory.mb	3584	3584
yarn.app.mapreduce.am.resource.mb	3584	3584
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m4.10xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1557m	-Xmx1557m
mapreduce.java.opts	-Xmx3114m	-Xmx3114m
mapreduce.map.memory.mb	1946	1946
mapreduce.reduce.memory.mb	3892	3892
yarn.app.mapreduce.am.resource.mb	3892	3892
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	155648	31104
yarn.nodemanager.resource.memory-mb	155648	31104

m4.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx1587m	-Xmx1587m
mapreduce.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.map.memory.mb	1 984	1 984
mapreduce.reduce.memory.mb	3 968	3 968
yarn.app.mapreduce.am.resource.mb	3 968	3 968

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253 952	31744
yarn.nodemanager.resource.memory-mb	253 952	31744

Instances m5

m5.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m5.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m5.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m5.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

m5.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

m5.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3 968	3 968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253 952	31744
yarn.nodemanager.resource.memory-mb	253 952	31744

m5.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

Instances m5a

m5a.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m5a.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m5a.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m5a.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

m5a.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

m5a.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3 968	3 968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253 952	31744
yarn.nodemanager.resource.memory-mb	253 952	31744

m5a.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

instances m5ad

m5ad.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m5ad.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5ad.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m5ad.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m5ad.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m5ad.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m5ad.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

Instances m5d

m5d.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m5d.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m5d.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m5d.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

m5d.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

m5d.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3 968	3 968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253 952	31744
yarn.nodemanager.resource.memory-mb	253 952	31744

m5d.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

instances m5dn

m5dn.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m5dn.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5dn.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m5dn.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m5dn.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m5dn.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m5dn.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

instances m5n

m5n.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m5n.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5n.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m5n.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m5n.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m5n.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m5n.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

instances m5zn

m5zn.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.java.opts	-Xmx4608m	-Xmx4608m
mapreduce.map.memory.mb	2880	2880
mapreduce.reduce.memory.mb	5760	5760
yarn.app.mapreduce.am.resource.mb	5760	5760
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

m5zn.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5zn.3xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2577m	-Xmx2577m
mapreduce.java.opts	-Xmx5154m	-Xmx5154m
mapreduce.map.memory.mb	3221	3221
mapreduce.reduce.memory.mb	6442	6442
yarn.app.mapreduce.am.resource.mb	6442	6442

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	38656	19328
yarn.nodemanager.resource.memory-mb	38656	19328

m5zn.6xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.java.opts	-Xmx5700m	-Xmx5700m
mapreduce.map.memory.mb	3563	3563
mapreduce.reduce.memory.mb	7126	7126
yarn.app.mapreduce.am.resource.mb	7126	7126
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	28496
yarn.nodemanager.resource.memory-mb	85504	28496

m5zn.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

Instances m6a

m6a.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6a.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6a.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6a.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6a.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m6a.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6a.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6a.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

m6a.48xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

Instances m6g

m6g.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6g.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6g.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6g.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6g.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6g.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

instances m6gd

m6gd.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6gd.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6gd.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6gd.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6gd.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6gd.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instances m6i

m6i.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6i.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6i.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6i.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6i.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6i.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6i.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6i.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instances m6id

m6id.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6id.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6id.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6id.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6id.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m6id.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6id.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6id.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

instances m6idn

m6idn.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6idn.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6idn.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6idn.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6idn.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m6idn.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6idn.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6idn.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

instances m6in

m6in.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6in.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6in.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6in.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6in.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6in.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6in.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6in.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instances m7a

m7a.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7a.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7a.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7a.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7a.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m7a.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m7a.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m7a.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

m7a.48xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

Instances m7g

m7g.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7g.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7g.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7g.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7g.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m7g.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instances m7gd

m7gd.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7gd.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7gd.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7gd.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7gd.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m7gd.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instances m7i

m7i.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7i.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7i.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7i.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7i.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m7i.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m7i.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m7i.48xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

Instances m7i-flex

m7i-flex.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7i-flex.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7i-flex.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7i-flex.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

M7i-Flex. 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m7i-flex. 16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

instances M8G

8 mg.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

8 kg x 2 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

8 kg x 4 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

8 g x 8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

8 kg x 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

8 g x 16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

8 g, 24 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

8 kg. 48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

instances m8gd

m8gd.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

8 mm. 2 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

8 mm x 4 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

8 mm x 8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

8 mm x 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

8 mm x 16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

8 mm x 24 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

8 mm x 48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

Instances p2

p2.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.java.opts	-Xmx21708m	-Xmx21708m
mapreduce.map.memory.mb	13568	13568
mapreduce.reduce.memory.mb	27136	27136
yarn.app.mapreduce.am.resource.mb	27136	27136
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

p2.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

p2.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx9267m	-Xmx9267m
mapreduce.java.opts	-Xmx18534m	-Xmx18534m
mapreduce.map.memory.mb	11584	11584
mapreduce.reduce.memory.mb	23168	23168
yarn.app.mapreduce.am.resource.mb	23168	23168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	23168
yarn.nodemanager.resource.memory-mb	741376	23168

Instances p3

p3.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

p3.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

p3.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

instances p4d

p4d.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-XM x 9302 m	-XM x 9302 m
mapreduce.java.opts	-XMx18604m	-XMx18604m
mapreduce.map.memory.mb	11627	11627
mapreduce.reduce.memory.mb	23254	23254
yarn.app.mapreduce.am.resource.mb	23254	23254
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1116160	23222
yarn.nodemanager.resource.memory-mb	1116160	23222

Instances p5

p5.48xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx8294m	-Xmx8294m
mapreduce.java.opts	-Xmx16588m	-Xmx16588m
mapreduce.map.memory.mb	10368	10368
mapreduce.reduce.memory.mb	20736	20736

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.app.mapreduce.am.resource.mb	20736	20736
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990656	20736
yarn.nodemanager.resource.memory-mb	1990656	20736

Instances r3

r3.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r3.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r3.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r3.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instances r4

r4.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r4.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r4.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r4.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r4.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instances r5

r5.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

r5.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

r5.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

r5.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253 952	31744
yarn.nodemanager.resource.memory-mb	253 952	31744

r5.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

r5.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

r5.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

Instances r5a

r5a.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

r5a.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

r5a.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

r5a.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253 952	31744
yarn.nodemanager.resource.memory-mb	253 952	31744

r5a.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

r5a.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

r5a.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

instances r5ad

r5ad.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5ad.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5ad.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5ad.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5ad.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5ad.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6246m	-Xmx6246m
mapreduce.java.opts	-Xmx12492m	-Xmx12492m
mapreduce.map.memory.mb	7808	7808
mapreduce.reduce.memory.mb	15616	15616
yarn.app.mapreduce.am.resource.mb	15616	15616
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	499712	31232
yarn.nodemanager.resource.memory-mb	499712	31232

r5ad.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

instances r5b

r5b.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5b.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5b.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5b.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5b.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5b.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5b.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

Instances r5d

r5d.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

r5d.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

r5d.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

r5d.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253 952	31744
yarn.nodemanager.resource.memory-mb	253 952	31744

r5d.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

r5d.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

r5d.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

instances r5dn

r5dn.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5dn.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5dn.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5dn.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5dn.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5dn.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5dn.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

Instances r5n

r5n.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5n.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5n.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5n.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5n.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5n.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5n.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

Instances r6a

r6a.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6a.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6a.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6a.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6a.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6a.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6a.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6a.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

r6a.48xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instances r6g

r6g.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6g.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6g.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6g.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6g.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6g.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instances r6gd

r6gd.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6gd.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6gd.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6gd.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6gd.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6gd.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instances r6i

r6i.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6i.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6i.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6i.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6i.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6i.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6i.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6i.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6029m	-Xmx6029m
mapreduce.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.map.memory.mb	7536	7536
mapreduce.reduce.memory.mb	15072	15072
yarn.app.mapreduce.am.resource.mb	15072	15072

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	30144
yarn.nodemanager.resource.memory-mb	964608	30144

Instances r6id

r6id.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6id.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6id.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6id.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6id.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6id.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6id.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6id.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

instances r6idn

r6idn.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6idn.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6idn.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6idn.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6idn.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6idn.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6idn.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6idn.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

Instances r6in

r6in.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6in.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6in.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6in.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6in.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6in.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6in.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6in.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

Instances r7a

r7a.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7a.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7a.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7a.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7a.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7a.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7a.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r7a.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

r7a.48xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instances r7g

r7g.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7g.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7g.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7g.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7g.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7g.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instances r7gd

r7gd.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7gd.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7gd.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7gd.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7gd.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7gd.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instances R7i

r7i.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7i.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7i.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7i.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7i.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7i.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7i.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r7i.48xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instances r7iz

r7iz.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7iz.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7iz.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7iz.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7iz.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7iz.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7iz.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6029m	-Xmx6029m
mapreduce.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.map.memory.mb	7536	7536
mapreduce.reduce.memory.mb	15072	15072
yarn.app.mapreduce.am.resource.mb	15072	15072

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	30144
yarn.nodemanager.resource.memory-mb	964608	30144

instances R8G

r8g.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

8 g x 2 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

8 g x 4 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

8 g x 8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

8 g, 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

8 g, 16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

8 g, 24 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

8 g, 48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instances r8gd

r8gd.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r8gd.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r8gd.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

8 g x 8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

8 g x 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

8 g x 16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

8 g x 24 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

8 g x 48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instances x1

x1.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.java.opts	-Xmx24116m	-Xmx24116m
mapreduce.map.memory.mb	15072	15072
mapreduce.reduce.memory.mb	30144	30144
yarn.app.mapreduce.am.resource.mb	30144	30144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	30144
yarn.nodemanager.resource.memory-mb	964608	30144

x1.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12109m	-Xmx12109m
mapreduce.java.opts	-Xmx24218m	-Xmx24218m
mapreduce.map.memory.mb	15136	15136
mapreduce.reduce.memory.mb	30272	30272
yarn.app.mapreduce.am.resource.mb	30272	30272

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1937408	30272
yarn.nodemanager.resource.memory-mb	1937408	30272

Instances x1e

x1e.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx22682m	-Xmx22682m
mapreduce.java.opts	-Xmx45364m	-Xmx45364m
mapreduce.map.memory.mb	28352	28352
mapreduce.reduce.memory.mb	56704	56704
yarn.app.mapreduce.am.resource.mb	56704	56704
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	113408	0
yarn.nodemanager.resource.memory-mb	113408	0

x1e.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx23501m	-Xmx23501m
mapreduce.java.opts	-Xmx47002m	-Xmx47002m
mapreduce.map.memory.mb	29376	29376
mapreduce.reduce.memory.mb	58752	58752
yarn.app.mapreduce.am.resource.mb	58752	58752
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	235008	0
yarn.nodemanager.resource.memory-mb	235008	0

x1e.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx23910m	-Xmx23910m
mapreduce.java.opts	-Xmx47820m	-Xmx47820m
mapreduce.map.memory.mb	29888	29888
mapreduce.reduce.memory.mb	59776	59776
yarn.app.mapreduce.am.resource.mb	59776	59776

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	478208	0
yarn.nodemanager.resource.memory-mb	478208	0

x1e.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx24115m	-Xmx24115m
mapreduce.java.opts	-Xmx48230m	-Xmx48230m
mapreduce.map.memory.mb	30144	30144
mapreduce.reduce.memory.mb	60288	60288
yarn.app.mapreduce.am.resource.mb	60288	60288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	0
yarn.nodemanager.resource.memory-mb	964608	0

x1e.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx24218m	-Xmx24218m
mapreduce.java.opts	-Xmx48436m	-Xmx48436m
mapreduce.map.memory.mb	30272	30272
mapreduce.reduce.memory.mb	60544	60544
yarn.app.mapreduce.am.resource.mb	60544	60544
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1937408	0
yarn.nodemanager.resource.memory-mb	1937408	0

x1e.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx24269m	-Xmx24269m
mapreduce.java.opts	-Xmx48538m	-Xmx48538m
mapreduce.map.memory.mb	30336	30336
mapreduce.reduce.memory.mb	60672	60672
yarn.app.mapreduce.am.resource.mb	60672	60672

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	3883008	0
yarn.nodemanager.resource.memory-mb	3883008	0

Instances x2gd

x2gd.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.java.opts	-Xmx21708m	-Xmx21708m
mapreduce.map.memory.mb	13568	13568
mapreduce.reduce.memory.mb	27136	27136
yarn.app.mapreduce.am.resource.mb	27136	27136
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

x2gd.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.java.opts	-Xmx23348m	-Xmx23348m
mapreduce.map.memory.mb	14592	14592
mapreduce.reduce.memory.mb	29184	29184
yarn.app.mapreduce.am.resource.mb	29184	29184
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

x2gd.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12083m	-Xmx12083m
mapreduce.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.map.memory.mb	15104	15104
mapreduce.reduce.memory.mb	30208	30208
yarn.app.mapreduce.am.resource.mb	30208	30208

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

x2gd.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

x2gd.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.java.opts	-Xmx24712m	-Xmx24712m
mapreduce.map.memory.mb	15445	15445
mapreduce.reduce.memory.mb	30890	30890
yarn.app.mapreduce.am.resource.mb	30890	30890
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30906
yarn.nodemanager.resource.memory-mb	741376	30906

x2gd.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.java.opts	-Xmx24780m	-Xmx24780m
mapreduce.map.memory.mb	15488	15488
mapreduce.reduce.memory.mb	30976	30976
yarn.app.mapreduce.am.resource.mb	30976	30976

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

instances x2idn

x2idn.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.java.opts	-Xmx24780m	-Xmx24780m
mapreduce.map.memory.mb	15488	15488
mapreduce.reduce.memory.mb	30976	30976
yarn.app.mapreduce.am.resource.mb	30976	30976
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

x2idn.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12425m	-Xmx12425m
mapreduce.java.opts	-Xmx24850m	-Xmx24850m
mapreduce.map.memory.mb	15531	15531
mapreduce.reduce.memory.mb	31062	31062
yarn.app.mapreduce.am.resource.mb	31062	31062
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31030
yarn.nodemanager.resource.memory-mb	1490944	31030

x2idn.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12442m	-Xmx12442m
mapreduce.java.opts	-Xmx24884m	-Xmx24884m
mapreduce.map.memory.mb	15552	15552
mapreduce.reduce.memory.mb	31104	31104
yarn.app.mapreduce.am.resource.mb	31104	31104

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990656	31104
yarn.nodemanager.resource.memory-mb	1990656	31104

Instances x2iedn

x2iedn.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx23347m	-Xmx23347m
mapreduce.java.opts	-Xmx46694m	-Xmx46694m
mapreduce.map.memory.mb	29184	29184
mapreduce.reduce.memory.mb	58368	58368
yarn.app.mapreduce.am.resource.mb	58368	58368
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	0
yarn.nodemanager.resource.memory-mb	116736	0

x2iedn.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.java.opts	-Xmx48332m	-Xmx48332m
mapreduce.map.memory.mb	30208	30208
mapreduce.reduce.memory.mb	60416	60416
yarn.app.mapreduce.am.resource.mb	60416	60416
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	0
yarn.nodemanager.resource.memory-mb	241664	0

x2iedn.4xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.java.opts	-Xmx49152m	-Xmx49152m
mapreduce.map.memory.mb	30720	30720
mapreduce.reduce.memory.mb	61440	61440
yarn.app.mapreduce.am.resource.mb	61440	61440

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	0
yarn.nodemanager.resource.memory-mb	491520	0

x2iedn.8xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx24781m	-Xmx24781m
mapreduce.java.opts	-Xmx49562m	-Xmx49562m
mapreduce.map.memory.mb	30976	30976
mapreduce.reduce.memory.mb	61952	61952
yarn.app.mapreduce.am.resource.mb	61952	61952
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	0
yarn.nodemanager.resource.memory-mb	991232	0

x2iedn.16xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx24883m	-Xmx24883m
mapreduce.java.opts	-Xmx49766m	-Xmx49766m
mapreduce.map.memory.mb	31104	31104
mapreduce.reduce.memory.mb	62208	62208
yarn.app.mapreduce.am.resource.mb	62208	62208
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990656	0
yarn.nodemanager.resource.memory-mb	1990656	0

x2iedn.24xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx24918m	-Xmx24918m
mapreduce.java.opts	-Xmx49836m	-Xmx49836m
mapreduce.map.memory.mb	31147	31147
mapreduce.reduce.memory.mb	62294	62294
yarn.app.mapreduce.am.resource.mb	62294	62294

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2990080	-32
yarn.nodemanager.resource.memory-mb	2990080	-32

x2iedn.32xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx24934m	-Xmx24934m
mapreduce.java.opts	-Xmx49868m	-Xmx49868m
mapreduce.map.memory.mb	31168	31168
mapreduce.reduce.memory.mb	62336	62336
yarn.app.mapreduce.am.resource.mb	62336	62336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	3989504	0
yarn.nodemanager.resource.memory-mb	3989504	0

instances x8g

x 8 g x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.java.opts	-Xmx21708m	-Xmx21708m
mapreduce.map.memory.mb	13568	13568
mapreduce.reduce.memory.mb	27136	27136
yarn.app.mapreduce.am.resource.mb	27136	27136
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

8 x g, 2 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.java.opts	-Xmx23348m	-Xmx23348m
mapreduce.map.memory.mb	14592	14592
mapreduce.reduce.memory.mb	29184	29184
yarn.app.mapreduce.am.resource.mb	29184	29184

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

8 x 8 x 4 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12083m	-Xmx12083m
mapreduce.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.map.memory.mb	15104	15104
mapreduce.reduce.memory.mb	30208	30208
yarn.app.mapreduce.am.resource.mb	30208	30208
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

8 x 8 x 8 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

8 x 12 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.java.opts	-Xmx24712m	-Xmx24712m
mapreduce.map.memory.mb	15445	15445
mapreduce.reduce.memory.mb	30890	30890
yarn.app.mapreduce.am.resource.mb	30890	30890

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30906
yarn.nodemanager.resource.memory-mb	741376	30906

8 x 16 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.java.opts	-Xmx24780m	-Xmx24780m
mapreduce.map.memory.mb	15488	15488
mapreduce.reduce.memory.mb	30976	30976
yarn.app.mapreduce.am.resource.mb	30976	30976
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

8 g x 24 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx12425m	-Xmx12425m
mapreduce.java.opts	-Xmx24850m	-Xmx24850m
mapreduce.map.memory.mb	15531	15531
mapreduce.reduce.memory.mb	31062	31062
yarn.app.mapreduce.am.resource.mb	31062	31062
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31030
yarn.nodemanager.resource.memory-mb	1490944	31030

8 g x 48 x large

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-XMx12458 m	-XMx12458 m
mapreduce.java.opts	-XM 24916 m	-XM 24916 m
mapreduce.map.memory.mb	15573	15573
mapreduce.reduce.memory.mb	31146	31146
yarn.app.mapreduce.am.resource.mb	31146	31146

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2990080	31210
yarn.nodemanager.resource.memory-mb	2990080	31210

instances z1d

z1d.xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

z1d.2xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

z1d.3xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6007m	-Xmx6007m
mapreduce.java.opts	-Xmx12014m	-Xmx12014m
mapreduce.map.memory.mb	7509	7509
mapreduce.reduce.memory.mb	15018	15018
yarn.app.mapreduce.am.resource.mb	15018	15018

Option de configuration	Valeur par défaut	Avec HBase installé
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

z1d.6xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6281m	-Xmx6281m
mapreduce.java.opts	-Xmx12562m	-Xmx12562m
mapreduce.map.memory.mb	7851	7851
mapreduce.reduce.memory.mb	15702	15702
yarn.app.mapreduce.am.resource.mb	15702	15702
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

z1d.12xlarge

Option de configuration	Valeur par défaut	Avec HBase installé
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

Paramètres de configuration de démon Hadoop

Les paramètres du démon Hadoop sont différents selon le type d' EC2 instance utilisé par un nœud de cluster. Les tableaux suivants répertorient les paramètres de configuration pour chaque type de l' EC2 instance.

Pour personnaliser ces paramètres, utilisez la classification de configuration `hadoop-env`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Types d'instances

- Instances c1
- Instances c3
- Instances c4
- Instances c5

- Instances c5a
- Instances c5ad
- Instances c5d
- Instances c5n
- Instances c6a
- instances c6g
- instances c6gd
- Instances c6gn
- Instances c6i
- Instances c6id
- Instances c6in
- instances C7a
- instances c7g
- Instances c7gd
- Instances c7gn
- Instances c7i
- Instances c7i-flex
- instances c8g
- instances c8gd
- Instances d2
- Instances d3
- instances d3en
- instances f2
- Instances g3
- Instances g3s
- Instances g4dn
- Instances g5
- instances g6g.6g.6
- Instances g6e
- Instances gr6

- Instances h1
- Instances i2
- Instances i3
- Instances i3en
- Instances i4g
- Instances i4i
- instances i7i
- instances i7ie
- instances i8g
- Instances im4gn
- Instances is4gen
- Instances m1
- Instances m2
- Instances m3
- Instances m4
- Instances m5
- Instances m5a
- instances m5ad
- Instances m5d
- instances m5dn
- instances m5n
- instances m5zn
- Instances m6a
- Instances m6g
- instances m6gd
- Instances m6i
- Instances m6id
- instances m6idn
- instances m6in
- Instances m7a

- Instances m7g
- Instances m7gd
- Instances m7i
- Instances m7i-flex
- instances M8G
- instances m8gd
- Instances p2
- Instances p3
- Instances P4d
- Instances p5
- Instances r3
- Instances r4
- Instances r5
- Instances r5a
- instances r5ad
- instances r5b
- Instances r5d
- instances r5dn
- Instances r5n
- Instances r6a
- instances r6g
- Instances r6gd
- Instances r6i
- Instances r6id
- instances r6idn
- Instances r6in
- Instances r7a
- instances r7g
- Instances r7gd
- Instances r7i

- Instances r7iz
- instances R8G
- instances r8gd
- Instances x1
- Instances x1e
- Instances x2gd
- instances x2idn
- Instances x2iedn
- instances x8g
- instances z1d

Instances c1

c1.medium

Paramètre	Valeur
YARN_RESOURCEMANAGER_HEAPSIZE	192
YARN_PROXYSERVER_HEAPSIZE	96
YARN_NODEMANAGER_HEAPSIZE	128
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	128
HADOOP_NAMENODE_HEAPSIZE	192
HADOOP_DATANODE_HEAPSIZE	96

c1.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAGER_HEAPSIZE	768

Paramètre	Valeur
YARN_PROXYSERVER_HEAPSIZE	384
YARN_NODEMANAGER_HEAPSIZE	512
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	512
HADOOP_NAMENODE_HEAPSIZE	768
HADOOP_DATANODE_HEAPSIZE	384

Instances c3

c3.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c3.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2396

Paramètre	Valeur
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

c3.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

c3.8.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317

Paramètre	Valeur
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

Instances c4

c4.large

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	1152
YARN_PROXYSERVER_HEAPSIZE	1152
YARN_NODEMANAGER_HEAPSIZE	1152
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1152
HADOOP_NAMENODE_HEAPSIZE	576
HADOOP_DATANODE_HEAPSIZE	384

c4.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124

Paramètre	Valeur
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c4.2.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

c4.4.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

c4.8.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

Instances c5

c5.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

c5.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

c5.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

c5.9xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3563
YARN_PROXYSERVER_HEAPSIZE	3563
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3563
HADOOP_NAMENODE_HEAPSIZE	7577
HADOOP_DATANODE_HEAPSIZE	1925

c5.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5.18xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5038
YARN_PROXYSERVER_HEAPSIZE	5038
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5038
HADOOP_NAMENODE_HEAPSIZE	14950
HADOOP_DATANODE_HEAPSIZE	3399

c5.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

Instances c5a

c5a.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c5a.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c5a.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c5a.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c5a.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5a.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c5a.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

Instances c5ad

c5ad.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c5ad.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c5ad.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c5ad.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c5ad.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c5ad.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c5ad.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

Instances c5d

c5d.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

c5d.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

c5d.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

c5d.9xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3563
YARN_PROXYSERVER_HEAPSIZE	3563
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3563
HADOOP_NAMENODE_HEAPSIZE	7577
HADOOP_DATANODE_HEAPSIZE	1925

c5d.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5d.18xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5038
YARN_PROXYSERVER_HEAPSIZE	5038
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5038
HADOOP_NAMENODE_HEAPSIZE	14950
HADOOP_DATANODE_HEAPSIZE	3399

c5d.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

Instances c5n

c5n.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2304
YARN_PROXYSERVER_HEAPSIZE	2304
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2304
HADOOP_NAMENODE_HEAPSIZE	1280
HADOOP_DATANODE_HEAPSIZE	665

c5n.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2519
YARN_PROXYSERVER_HEAPSIZE	2519
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2519
HADOOP_NAMENODE_HEAPSIZE	2355
HADOOP_DATANODE_HEAPSIZE	880

c5n.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2949
YARN_PROXYSERVER_HEAPSIZE	2949
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2949
HADOOP_NAMENODE_HEAPSIZE	4505
HADOOP_DATANODE_HEAPSIZE	1310

c5n.9xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5n.18xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

Instances c6a

c6a.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6a.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6a.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6a.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6a.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6a.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6a.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6a.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

c6a.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

instances c6g

c6g.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6g.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6g.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6g.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6g.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6g.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

instances c6gd

c6gd.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6gd.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6gd.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6gd.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6gd.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6gd.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instances c6gn

c6gn.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6gn.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6gn.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6gn.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6gn.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6gn.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instances c6i

c6i.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6i.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6i.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6i.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6i.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6i.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6i.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6i.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instances c6id

c6id.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6id.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6id.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6id.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6id.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6id.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6id.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6id.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instances c6in

c6in.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6in.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6in.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6in.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6in.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6in.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6in.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6in.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

instances C7a

c7a.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7a.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7a.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7a.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7a.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7a.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c7a.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c7a.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

c7a.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

instances c7g

c7g.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7g.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7g.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7g.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7g.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7g.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instances c7gd

c7gd.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7gd.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7gd.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7gd.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7gd.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7gd.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instances c7gn

c7gn.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7gn.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7gn.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7gn.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7gn.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7gn.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instances c7i

c7i.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7i.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7i.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7i.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7i.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7i.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c7i.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c7i.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instances c7i-flex

c7i-flex.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7i-flex.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7i-flex.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7i-flex.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7i-flex.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7i-flex.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

instances c8g

c8g.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c8g.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c8g.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c8g.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c8g.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c8g.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c8g.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c8g.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

instances c8gd

c8gd.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c8gd.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c8gd.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c8gd.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c8gd.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c8gd.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c8gd.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c8g x 48 x large

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instances d2

d2.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

d2.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

d2.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

d2.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instances d3

d3.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

d3.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

d3.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

d3.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

instances d3en

d3en.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

d3en.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

d3en.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

d3en.6xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

d3en.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

d3en.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

instances f2

F2,6 x large

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

f2.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

f2.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

Instances g3

g3.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

g3.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

g3.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instances g3s

g3s.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

Instances g4dn

g4dn.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416

Paramètre	Valeur
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

g4dn.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

g4dn.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

g4dn.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

g4dn.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

g4dn.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

Instances g5

g5.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

g5.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

g5.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

g5.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

g5.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

g5.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

g5.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

g5.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

instances g6g.6g.6

g6.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

g6.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

g6.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

g6.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4567
YARN_PROXYSERVER_HEAPSIZE	4567
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4567
HADOOP_NAMENODE_HEAPSIZE	12595
HADOOP_DATANODE_HEAPSIZE	2928

g6.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5795
YARN_PROXYSERVER_HEAPSIZE	5795
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5795
HADOOP_NAMENODE_HEAPSIZE	18739
HADOOP_DATANODE_HEAPSIZE	4096

g6.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7045
YARN_PROXYSERVER_HEAPSIZE	7045
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7045
HADOOP_NAMENODE_HEAPSIZE	24985
HADOOP_DATANODE_HEAPSIZE	4096

g6.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9543
YARN_PROXYSERVER_HEAPSIZE	9543
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9543
HADOOP_NAMENODE_HEAPSIZE	37478
HADOOP_DATANODE_HEAPSIZE	4096

g6.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	16998
YARN_PROXYSERVER_HEAPSIZE	16998
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	16998
HADOOP_NAMENODE_HEAPSIZE	74752
HADOOP_DATANODE_HEAPSIZE	4096

Instances g6e

g6e.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

g6e.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

g6e.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

g6e.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

g6e.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

g6e.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

g6e.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

g6e.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instances gr6

gr6.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4567
YARN_PROXYSERVER_HEAPSIZE	4567
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4567
HADOOP_NAMENODE_HEAPSIZE	12595
HADOOP_DATANODE_HEAPSIZE	2928

gr6.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7045
YARN_PROXYSERVER_HEAPSIZE	7045
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7045
HADOOP_NAMENODE_HEAPSIZE	24985
HADOOP_DATANODE_HEAPSIZE	4096

Instances h1

h1.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

h1.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

h1.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

h1.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

Instances i2

i2.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i2.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i2.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i2.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instances i3

i3.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i3.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i3.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i3.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i3.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instances i3en

i3en.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

i3en.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

i3en.3xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

i3en.6xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

i3en.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

i3en.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

Instances i4g

i4g.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i4g.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i4g.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i4g.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i4g.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instances i4i

i4i.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i4i.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i4i.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i4i.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i4i.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i4i.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i4i.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i4i.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

instances i7i

i7i.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i7i.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i7i.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i7i.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i7i.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i7i.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i7i.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i7i.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

instances i7ie

i7ie.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i7ie.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i7ie.3xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

i7ie.6xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.18xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	13332
YARN_PROXYSERVER_HEAPSIZE	13332
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	13332
HADOOP_NAMENODE_HEAPSIZE	56422
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

instances i8g

i8g.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i8g.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i8g.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i8g.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i8g.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i8g.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i8g.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i8g.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instances im4gn

im4gn.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

im4gn.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

im4gn.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

im4gn.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

im4gn.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instances is4gen

is4gen.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2557
YARN_PROXYSERVER_HEAPSIZE	2557
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2557
HADOOP_NAMENODE_HEAPSIZE	2547
HADOOP_DATANODE_HEAPSIZE	919

is4gen.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3025
YARN_PROXYSERVER_HEAPSIZE	3025
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3025
HADOOP_NAMENODE_HEAPSIZE	4889
HADOOP_DATANODE_HEAPSIZE	1387

is4gen.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

is4gen.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

Instances m1

m1.small

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	256
YARN_PROXYSERVER_HEAPSIZE	96
YARN_NODEMANAGER_HEAPSIZE	192
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	128
HADOOP_NAMENODE_HEAPSIZE	192
HADOOP_DATANODE_HEAPSIZE	96

m1.medium

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	384
YARN_PROXYSERVER_HEAPSIZE	192
YARN_NODEMANAGER_HEAPSIZE	256

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	256
HADOOP_NAMENODE_HEAPSIZE	384
HADOOP_DATANODE_HEAPSIZE	192

m1.large

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	768
YARN_PROXYSERVER_HEAPSIZE	384
YARN_NODEMANAGER_HEAPSIZE	512
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	512
HADOOP_NAMENODE_HEAPSIZE	768
HADOOP_DATANODE_HEAPSIZE	384

m1.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	1024
YARN_PROXYSERVER_HEAPSIZE	512
YARN_NODEMANAGER_HEAPSIZE	768

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1024
HADOOP_NAMENODE_HEAPSIZE	2304
HADOOP_DATANODE_HEAPSIZE	384

Instances m2

m2.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	1536
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1024
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1024
HADOOP_NAMENODE_HEAPSIZE	3072
HADOOP_DATANODE_HEAPSIZE	384

m2.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	1536
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1024

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1536
HADOOP_NAMENODE_HEAPSIZE	6144
HADOOP_DATANODE_HEAPSIZE	384

m2.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2048
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1536
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1536
HADOOP_NAMENODE_HEAPSIZE	12288
HADOOP_DATANODE_HEAPSIZE	384

Instances m3

m3.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

m3.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

Instances m4

m4.large

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

m4.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m4.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m4.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m4.10xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5365
YARN_PROXYSERVER_HEAPSIZE	5365
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5365
HADOOP_NAMENODE_HEAPSIZE	16588
HADOOP_DATANODE_HEAPSIZE	3727

m4.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

Instances m5

m5.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m5.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m5.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m5.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

m5.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

m5.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

Instances m5a

m5a.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m5a.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m5a.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m5a.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

m5a.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

m5a.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5a.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

instances m5ad

m5ad.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m5ad.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5ad.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m5ad.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m5ad.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m5ad.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m5ad.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instances m5d

m5d.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m5d.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m5d.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m5d.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

m5d.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

m5d.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5d.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

instances m5dn

m5dn.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m5dn.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5dn.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m5dn.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m5dn.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m5dn.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m5dn.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

instances m5n

m5n.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m5n.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5n.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m5n.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m5n.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m5n.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m5n.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

instances m5zn

m5zn.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

m5zn.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5zn.3xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3025
YARN_PROXYSERVER_HEAPSIZE	3025
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3025
HADOOP_NAMENODE_HEAPSIZE	4889
HADOOP_DATANODE_HEAPSIZE	1387

m5zn.6xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

m5zn.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

Instances m6a

m6a.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6a.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6a.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6a.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6a.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6a.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6a.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6a.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

m6a.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instances m6g

m6g.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6g.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6g.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6g.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6g.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6g.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

instances m6gd

m6gd.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6gd.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6gd.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6gd.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6gd.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6gd.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instances m6i

m6i.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6i.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6i.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6i.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6i.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6i.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6i.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6i.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instances m6id

m6id.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6id.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6id.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6id.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6id.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6id.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6id.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6id.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

instances m6idn

m6idn.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6idn.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6idn.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6idn.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6idn.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6idn.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6idn.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6idn.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

instances m6in

m6in.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6in.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6in.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6in.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6in.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6in.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6in.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6in.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instances m7a

m7a.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7a.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7a.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7a.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7a.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7a.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m7a.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m7a.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

m7a.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instances m7g

m7g.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7g.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7g.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7g.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7g.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7g.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instances m7gd

m7gd.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7gd.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7gd.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7gd.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7gd.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7gd.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instances m7i

m7i.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7i.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7i.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7i.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7i.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m7i.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m7i.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m7i.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instances m7i-flex

m7i-flex.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7i-flex.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7i-flex.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7i-flex.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7i-flex.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7i-flex.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

instances M8G

m8g.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m8g.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m8g.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m8g.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m8g.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m8g.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m8g.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m8g.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

instances m8gd

m8gd.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m8gd.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m8gd.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m8gd.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m8gd.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m8gd.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m8gd.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

8 mm x 48 x large

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instances p2

p2.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

p2.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

p2.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instances p3

p3.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

p3.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

p3.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instances P4d

p4d.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	24576
YARN_PROXYSERVER_HEAPSIZE	24576
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	24576
HADOOP_NAMENODE_HEAPSIZE	112640
HADOOP_DATANODE_HEAPSIZE	4096

Instances p5

p5.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065

Paramètre	Valeur
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

Instances r3

r3.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r3.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338

Paramètre	Valeur
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r3.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r3.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instances r4

r4.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r4.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r4.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r4.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r4.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instances r5

r5.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

r5.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

r5.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

r5.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

r5.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

r5.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

r5.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

Instances r5a

r5a.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

r5a.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

r5a.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

r5a.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

r5a.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

r5a.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

r5a.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

instances r5ad

r5ad.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5ad.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5ad.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5ad.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5ad.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5ad.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12247
YARN_PROXYSERVER_HEAPSIZE	12247
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12247
HADOOP_NAMENODE_HEAPSIZE	50995
HADOOP_DATANODE_HEAPSIZE	4096

r5ad.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

instances r5b

r5b.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5b.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5b.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5b.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5b.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5b.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5b.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instances r5d

r5d.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

r5d.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

r5d.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

r5d.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

r5d.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

r5d.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

r5d.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

instances r5dn

r5dn.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5dn.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5dn.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5dn.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5dn.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5dn.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5dn.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instances r5n

r5n.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5n.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5n.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5n.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5n.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5n.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5n.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instances r6a

r6a.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6a.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6a.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6a.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6a.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6a.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6a.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6a.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

r6a.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

instances r6g

r6g.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6g.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6g.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6g.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6g.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6g.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instances r6gd

r6gd.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6gd.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6gd.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6gd.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6gd.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6gd.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instances r6i

r6i.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6i.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6i.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6i.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6i.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6i.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6i.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6i.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

Instances r6id

r6id.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6id.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6id.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6id.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6id.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6id.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6id.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6id.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

instances r6idn

r6idn.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6idn.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6idn.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6idn.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

Instances r6in

r6in.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6in.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6in.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6in.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6in.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6in.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6in.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6in.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

Instances r7a

r7a.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7a.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7a.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7a.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7a.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7a.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7a.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r7a.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

r7a.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

instances r7g

r7g.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7g.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7g.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7g.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7g.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7g.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instances r7gd

r7gd.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7gd.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7gd.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7gd.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7gd.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7gd.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instances r7i

r7i.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7i.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7i.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7i.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7i.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7i.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7i.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r7i.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instances r7iz

r7iz.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7iz.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7iz.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7iz.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7iz.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7iz.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7iz.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

instances R8G

r8g.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r8g.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r8g.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r8g.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r8g.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r8g.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r8g.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r8g.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

instances r8gd

r8gd.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r8gd.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r8gd.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r8gd.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

8 g x 48 x large

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instances x1

x1.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

x1.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	41000
YARN_PROXYSERVER_HEAPSIZE	41000
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	41000
HADOOP_NAMENODE_HEAPSIZE	194764
HADOOP_DATANODE_HEAPSIZE	4096

Instances x1e

x1e.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4520
YARN_PROXYSERVER_HEAPSIZE	4520
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4520
HADOOP_NAMENODE_HEAPSIZE	12364
HADOOP_DATANODE_HEAPSIZE	2882

x1e.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	6952
YARN_PROXYSERVER_HEAPSIZE	6952
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6952
HADOOP_NAMENODE_HEAPSIZE	24524
HADOOP_DATANODE_HEAPSIZE	4096

x1e.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	11816
YARN_PROXYSERVER_HEAPSIZE	11816
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	11816
HADOOP_NAMENODE_HEAPSIZE	48844
HADOOP_DATANODE_HEAPSIZE	4096

x1e.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

x1e.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	41000
YARN_PROXYSERVER_HEAPSIZE	41000
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	41000
HADOOP_NAMENODE_HEAPSIZE	194764
HADOOP_DATANODE_HEAPSIZE	4096

x1e.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	79912
YARN_PROXYSERVER_HEAPSIZE	79912
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	79912
HADOOP_NAMENODE_HEAPSIZE	389324
HADOOP_DATANODE_HEAPSIZE	4096

Instances x2gd

x2gd.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

x2gd.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

x2gd.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

x2gd.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

x2gd.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

x2gd.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

instances x2idn

x2idn.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x2idn.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

x2idn.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

Instances x2iedn

x2iedn.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

x2iedn.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	62054
YARN_PROXYSERVER_HEAPSIZE	62054
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	62054
HADOOP_NAMENODE_HEAPSIZE	300032
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.32xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	82042
YARN_PROXYSERVER_HEAPSIZE	82042
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	82042
HADOOP_NAMENODE_HEAPSIZE	399974
HADOOP_DATANODE_HEAPSIZE	4096

instances x8g

x8g.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

x8g.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

x8g.4xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

x8g.8xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

x8g.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

x8g.16xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x8g.24xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

x8g.48xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	62054
YARN_PROXYSERVER_HEAPSIZE	62054
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	62054
HADOOP_NAMENODE_HEAPSIZE	300032
HADOOP_DATANODE_HEAPSIZE	4096

instances z1d

z1d.xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

z1d.2xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

z1d.3xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

z1d.6xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

z1d.12xlarge

Paramètre	Valeur
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048

Paramètre	Valeur
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

Configuration HDFS

Le tableau suivant décrit les paramètres par défaut du système de fichiers distribué Hadoop (HDFS) et leurs configurations. Vous pouvez modifier ces valeurs à l'aide de la classification de configuration `hdfs-site`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Warning

1. Paramétrer `dfs.replication` sur la valeur 1 avec les clusters de moins de quatre nœuds peut entraîner une perte de données HDFS en cas de panne d'un seul nœud. Si votre cluster dispose d'un stockage HDFS, nous vous recommandons de le configurer avec au moins quatre nœuds principaux pour les charges de travail de production afin d'éviter toute perte de données.
2. Amazon EMR n'autorisera pas les clusters à dimensionner les nœuds primaires situés en dessous.`dfs.replication` Par exemple, si `dfs.replication = 2`, le nombre minimum de nœuds principaux est 2.
3. Lorsque vous utilisez la mise à l'échelle gérée, autoscaling, ou que vous choisissez de redimensionner manuellement votre cluster, nous vous recommandons de définir `dfs.replication` sur une valeur supérieure ou égale à 2.

Paramètre	Définition	Valeur par défaut
`dfs.block.size`	Taille des blocs HDFS. Lorsque vous exploitez des données stockées dans HDFS, la taille de division est généralement celle d'un bloc HDFS. Plus la taille est grande, plus la granularité de la tâche est faible. En	134217728 (128 Mo)

Paramètre	Définition	Valeur par défaut
	revanche, les contraintes sur le `NameNode` du cluster sont moins importantes.	
`dfs.replication`	Le nombre de copies de chaque bloc à stocker pour des questions de durabilité. Amazon EMR définit cette valeur en fonction du nombre de nœuds principaux avec lesquels le cluster est approvisionné. Ajustez la valeur en fonction de vos besoins. Pour remplacer la valeur par défaut, utilisez la classification `hdfs-site` .	1 pour les clusters provisionnés avec moins de quatre nœuds principaux 2 pour les clusters approvisionnés avec moins de dix nœuds principaux 3 pour tous les autres clusters

Chiffrement transparent dans HDFS sur Amazon EMR

Le chiffrement transparent est mis en œuvre grâce à l'utilisation des zones de chiffrement HDFS qui sont des chemins d'accès HDFS que vous définissez. Chaque zone de chiffrement a sa propre clé, qui est stockée dans le serveur de clés spécifié à l'aide de la classification de configuration `hdfs-site`.

A partir de la version 4.8.0 d'Amazon EMR, vous pouvez utiliser les configurations de sécurité Amazon EMR pour configurer plus facilement les paramètres de chiffrement des données pour les clusters. Les configurations de sécurité offrent des paramètres permettant d'activer la sécurité des données en transit et des données au repos dans les volumes de stockage Amazon Elastic Block Store (Amazon EBS) et les données EMRFS dans Amazon S3. Pour plus d'informations, consultez Chiffrement des données en transit et au repos dans le Guide de gestion d'Amazon EMR.

Amazon EMR utilise le KMS Hadoop par défaut ; vous pouvez toutefois utiliser un autre KMS qui implémente l'opération d'API. KeyProvider Chaque fichier dans une zone de chiffrement HDFS a sa propre clé de chiffrement de données unique qui est chiffrée par la clé de la zone de chiffrement. Les données HDFS sont chiffrées end-to-end (au repos et en transit) lorsque les données sont écrites

dans une zone de chiffrement, car les activités de chiffrement et de déchiffrement ne se produisent que dans le client.

Vous ne pouvez pas déplacer des fichiers entre les zones de chiffrement ou d'une zone de chiffrement vers des chemins d'accès non chiffrés.

Le client HDFS NameNode et Hadoop interagissent avec le KMS Hadoop (ou un autre KMS que vous avez configuré) via l' KeyProvider opération d'API. Le KMS est responsable du stockage des clés de chiffrement dans le magasins de clés de sauvegarde. En outre, Amazon EMR inclut la stratégie force illimitée JCE afin que vous puissiez créer des clés à la longueur souhaitée.

Pour plus d'informations, consultez Chiffrement transparent dans HDFS dans la documentation Hadoop.

> **Note**
>
> Dans Amazon EMR, KMS sur HTTPS n'est pas activé par défaut avec Hadoop KMS. Pour plus d'informations sur l'activation de KMS sur HTTPS, consultez la Documentation Hadoop KMS.

Configuration du chiffrement transparent HDFS

Vous pouvez configurer le chiffrement transparent dans Amazon EMR en créant des clés et en ajoutant des zones de chiffrement. Pour ce faire, plusieurs options s'offrent à vous :

- Avec l'opération API de configuration Amazon EMR lorsque vous créez un cluster
- Avec une étape Hadoop JAR avec command-runner.jar
- En vous connectant au nœud principal du cluster Hadoop et en utilisant les clients de ligne de commande `hadoop key` et `hdfs crypto`
- Utilisation du REST APIs pour Hadoop KMS et HDFS

Pour plus d'informations sur le REST APIs, consultez la documentation correspondante pour Hadoop KMS et HDFS.

Pour créer des zones de chiffrement et leurs clés lors de la création du cluster à l'aide de l'interface de ligne de commande

La classification `hdfs-encryption-zones` dans l'opération d'API de configuration vous permet de spécifier un nom de clé et une zone de chiffrement lorsque vous créez un cluster. Amazon EMR crée cette clé dans Hadoop KMS sur votre cluster et configure la zone de chiffrement.

- Créez un cluster à l'aide de la commande suivante.

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

`myConfig.json`:

```
[
 {
  "Classification": "hdfs-encryption-zones",
  "Properties": {
   "/myHDFSPath1": "path1_key",
   "/myHDFSPath2": "path2_key"
  }
 }
]
```

Pour créer des zones de chiffrement et leurs clés manuellement sur le nœud maître

1. Lancement de votre cluster à l'aide d'une version Amazon EMR supérieure à 4.1.0.
2. Connectez-vous au nœud principal du cluster avec SSH.
3. Créez une clé au sein d'Hadoop KMS.

```
$ hadoop key create path2_key
path2_key has been successfully created with options Options{cipher='AES/CTR/
NoPadding', bitLength=256, description='null', attributes=null}.
KMSClientProvider[http://ip-x-x-x-x.ec2.internal:16000/kms/v1/] has been updated.
```

> **Important**
>
> Hadoop KMS a besoin que vos noms de clés soient en minuscules. Si vous utilisez une clé comprenant des majuscules, votre cluster échouera lors du lancement.

4. Créez le chemin d'accès de zone de chiffrement dans HDFS.

```
$ hadoop fs -mkdir /myHDFSPath2
```

5. Faites du chemin d'accès HDFS une zone de chiffrement à l'aide de la clé que vous avez créée.

```
$ hdfs crypto -createZone -keyName path2_key -path /myHDFSPath2
Added encryption zone /myHDFSPath2
```

Pour créer des zones de chiffrement et leurs clés manuellement à l'aide du AWS CLI

- Ajoutez des étapes pour créer les clés KMS et les zones de chiffrement manuellement avec la commande suivante.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps Type=CUSTOM_JAR,Name="Create
 First Hadoop KMS Key",Jar="command-runner.jar",ActionOnFailure=CONTINUE,Args=[/
bin/bash,-c,"\"hadoop key create path1_key\""] \
Type=CUSTOM_JAR,Name="Create First Hadoop HDFS Path",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop fs -mkdir /
myHDFSPath1\""] \
Type=CUSTOM_JAR,Name="Create First Encryption Zone",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hdfs crypto -createZone
 -keyName path1_key -path /myHDFSPath1\""] \
Type=CUSTOM_JAR,Name="Create Second Hadoop KMS Key",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop key create
 path2_key\""] \
Type=CUSTOM_JAR,Name="Create Second Hadoop HDFS Path",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop fs -mkdir /
myHDFSPath2\""] \
```

```
Type=CUSTOM_JAR,Name="Create Second Encryption Zone",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hdfs crypto -createZone
 -keyName path2_key -path /myHDFSPath2\""]
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

Considérations pour le chiffrement transparent HDFS

Une bonne pratique consiste à créer une zone de chiffrement pour chaque application où des fichiers peuvent être écrits. Vous pouvez également chiffrer l'ensemble du HDFS en utilisant la hdfs-encryption-zones classification dans l'API de configuration et en spécifiant le chemin racine (/) comme zone de chiffrement.

Serveur de gestion des clés Hadoop

Hadoop KMS est un serveur de gestion des clés qui offre la possibilité de mettre en place des services de chiffrement pour les clusters Hadoop et peut servir de fournisseur de clé pour Chiffrement transparent dans HDFS sur Amazon EMR. Hadoop KMS dans Amazon EMR est installé et activé par défaut lorsque vous sélectionnez l'application Hadoop lors du lancement d'un cluster EMR. Le KMS Hadoop ne stocke pas les clés lui-même, sauf dans le cas de mise en cache temporaire. Hadoop KMS agit comme un proxy entre le fournisseur de clés et le tiers de confiance du client dans un magasins de clés de sauvegarde. Il ne s'agit pas d'un magasin de clés. Le keystore par défaut créé pour Hadoop KMS est l'extension KeyStore de cryptographie Java (JCEKS). La stratégie de force illimitée JCE est également incluse, vous pouvez donc créer des clés avec la longueur souhaitée. Hadoop KMS prend également en charge une gamme de solutions ACLs qui contrôlent l'accès aux clés et aux opérations clés indépendamment des autres applications clientes telles que HDFS. La longueur de clé par défaut dans Amazon EMR est de 256 bits.

Pour configurer Hadoop KMS, utilisez la hadoop-kms-site classification pour modifier les paramètres. Pour configurer ACLs, vous utilisez la classification kms-acls.

Pour plus d'informations, consultez la Documentation Hadoop KMS. Hadoop KMS est utilisé dans le chiffrement transparent Hadoop HDFS. Pour en savoir plus sur le chiffrement transparent HDFS, consultez Chiffrement transparent HDFS dans la documentation Apache Hadoop.

> **Note**
>
> Dans Amazon EMR, KMS sur HTTPS n'est pas activé par défaut avec Hadoop KMS. Pour savoir comment activer KMS sur HTTPS, consultez la Documentation KMS Hadoop.

> **Important**
>
> Hadoop KMS a besoin que vos noms de clés soient en minuscules. Si vous utilisez une clé comprenant des majuscules, votre cluster échouera lors du lancement.

Configuration de Hadoop KMS dans Amazon EMR

Avec Amazon EMR version 4.6.0 ou ultérieure, `kms-http-port` est 9700 et `kms-admin-port` est 9701.

Vous pouvez configurer Hadoop KMS au moment de la création du cluster à l'aide de l'API de configuration pour versions Amazon EMR. Les éléments suivants correspondent aux classifications d'objet de configuration disponibles pour Hadoop KMS :

Classifications de configuration Hadoop KMS

Classification	Nom de fichier
hadoop-kms-site	`kms-site.xml`
hadoop-kms-acls	`kms-acls.xml`
hadoop-kms-env	`kms-env.sh`
hadoop-kms-log4j	`kms-log4j.properties`

Pour configurer Hadoop KMS à l' ACLs aide de la CLI

- Créez un cluster avec Hadoop KMS à ACLs l'aide de la commande suivante :

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
```

```
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

`myConfig.json`:

```
[
 {
  "Classification": "hadoop-kms-acls",
  "Properties": {
   "hadoop.kms.blacklist.CREATE": "hdfs,foo,myBannedUser",
   "hadoop.kms.acl.ROLLOVER": "myAllowedUser"
  }
 }
]
```

Pour désactiver le cache Hadoop KMS à l'aide de l'interface de ligne de commande

- Créez un cluster avec le paramètre Hadoop KMS `hadoop.kms.cache.enable` défini sur `false` à l'aide de la commande suivante :

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

`myConfig.json`:

```
[
 {
  "Classification": "hadoop-kms-site",
  "Properties": {
   "hadoop.kms.cache.enable": "false"
  }
 }
]
```

Pour définir des variables d'environnement le script **`kms-env.sh`** à l'aide de l'interface de ligne de commande

- Modifiez les paramètres dans `kms-env.sh` via la configuration `hadoop-kms-env`. Créez un cluster avec Hadoop KMS à l'aide de la commande suivante :

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

`myConfig.json`:

```
[
 {
  "Classification": "hadoop-kms-env",
  "Properties": {},
  "Configurations": [
   {
    "Classification": "export",
```

```
    "Properties": {
     "JAVA_LIBRARY_PATH": "/path/to/files",
     "KMS_SSL_KEYSTORE_FILE": "/non/Default/Path/.keystore",
     "KMS_SSL_KEYSTORE_PASS": "myPass"
    },
    "Configurations": []
   }
  ]
 }
]
```

Pour plus d'informations sur la configuration de Hadoop KMS, consultez la Documentation Hadoop KMS.

Chiffrement transparent HDFS sur les clusters EMR avec plusieurs nœuds principaux

Apache Ranger KMS est utilisé dans un cluster Amazon EMR avec plusieurs nœuds primaires pour un chiffrement transparent dans HDFS.

Apache Ranger KMS stocke sa clé racine et ses clés de zone de chiffrement (EZ) dans votre Amazon RDS pour un cluster Amazon EMR comportant plusieurs nœuds primaires. Pour activer le chiffrement transparent dans HDFS sur un cluster Amazon EMR avec plusieurs nœuds primaires, vous devez fournir les configurations suivantes.

- Amazon RDS ou votre propre URL de connexion au serveur MySQL pour stocker la clé racine Ranger KMS et la clé EZ
- Nom d'utilisateur et mot de passe pour MySQL
- Mot de passe pour la clé racine KMS de Ranger
- Fichier PEM de l'autorité de certification (CA) pour la connexion SSL au serveur MySQL. Vous pouvez télécharger le bundle de certificats correspondant Région AWS à votre adresse Download certificate bundles for Amazon RDS.

Vous pouvez fournir ces configurations à l'aide des classifications `ranger-kms-dbks-site` et `ranger-kms-db-ca`, comme l'illustre l'exemple suivant.

```
[
```

```
 {
  "Classification": "ranger-kms-dbks-site",
  "Properties": {
   "ranger.ks.jpa.jdbc.url": "jdbc:log4jdbc:mysql://mysql-host-url.xx-
xxx-1.xxx.amazonaws.com:3306/rangerkms",
   "ranger.ks.jpa.jdbc.user": "mysql-user-name",
   "ranger.ks.jpa.jdbc.password": "mysql-password",
   "ranger.db.encrypt.key.password": "password-for-encrypting-a-master-key"
  }
 },
 {
  "Classification": "ranger-kms-db-ca",
  "Properties": {
   "ranger.kms.trust.ca.file.s3.url": "<S3-path-of-downloaded-pem-file>"
  }
 }
]
```

Voici les classifications d'objets de configuration pour Apache Ranger KMS.

Classifications de configuration Hadoop KMS

Classification	Description
ranger-kms-dbks-site	Modifiez les valeurs dans le fichier dbks-site .xml de Ranger KMS.
ranger-kms-site	Modifiez les valeurs dans le fichier ranger-kms-site .xml de Ranger KMS.
ranger-kms-env	Modifiez les valeurs dans l'environnement Ranger KMS.
ranger-kms-log4j	Modifiez les valeurs dans le fichier kms-log4j .properties de Ranger KMS.
ranger-kms-db-ca	Modifiez les valeurs du fichier CA sur S3 pour la connexion SSL MySQL avec Ranger KMS.

Considérations

- Il est vivement recommandé de chiffrer votre instance Amazon RDS pour améliorer la sécurité. Pour plus d'informations, consultez Présentation du chiffrement des ressources Amazon RDS.
- Il est fortement recommandé d'utiliser une base de données MySQL distincte pour chaque cluster Amazon EMR avec plusieurs nœuds primaires pour garantir un niveau de sécurité élevé.
- Pour configurer le chiffrement transparent dans HDFS sur un cluster Amazon EMR avec plusieurs nœuds primaires, vous devez spécifier la classification `hdfs-encryption-zones` lors de la création du cluster. Sinon, Ranger KMS ne sera pas configuré ou ne démarrera pas. La reconfiguration de la classification `hdfs-encryption-zones` ou de l'une des classifications de la configuration KMS de Hadoop sur un cluster en cours d'exécution n'est pas prise en charge sur un cluster Amazon EMR avec plusieurs nœuds primaires.
- Le bundle de certificats PEM que vous téléchargez depuis Download certificate bundles for Amazon RDS regroupe plusieurs certificats dans un seul fichier. Amazon EMR 7.3.0 et versions ultérieures prennent en charge l'importation de plusieurs certificats depuis le fichier PEM avec la configuration. `ranger.kms.trust.ca.file.s3.url`

Création ou exécution d'une application Hadoop

Rubriques

- Créer des fichiers binaires à l'aide d'Amazon EMR
- Traitement des données avec Streaming
- Traitement de données à l'aide d'un fichier JAR personnalisé

Créer des fichiers binaires à l'aide d'Amazon EMR

Vous pouvez utiliser Amazon EMR comme un environnement de génération pour compiler les programmes à utiliser dans votre cluster. Les programmes que vous utilisez avec Amazon EMR doivent être compilés sur un système exécutant la même version de Linux que celle utilisée par Amazon EMR. Pour une version 32 bits, les options de compilation sur un ordinateur 32 bits ou de compilation croisée 32 bits doivent être activées. Pour une version 64 bits, les options de compilation sur un ordinateur 64 bits ou de compilation croisée 64 bits doivent être activées. Pour plus d'informations sur les versions des EC2 instances, consultez Planifier et configurer EC2 les instances dans le guide de gestion Amazon EMR. Les langages de programmation pris en charge incluent C++, Python et C#.

Le tableau suivant décrit les étapes nécessaires pour générer et tester votre application à l'aide d'Amazon EMR.

Processus de création d'un module

1	Connectez-vous au nœud maître de votre cluster.
2	Copiez les fichiers source dans le nœud principal.
3	Créez des fichiers binaires avec les optimisations nécessaires.
4	Copiez des fichiers binaires à partir du nœud principal dans Amazon S3.

Chacune de ces étapes est détaillée dans les sections suivantes.

Pour vous connecter au nœud principal du cluster

- Suivez les instructions fournies dans Connexion au nœud principal à l'aide de SSH dans le Guide de gestion d'Amazon EMR.

Pour copier les fichiers source dans le nœud principal

1. Placez vos fichiers source dans un compartiment Amazon S3. Pour savoir comment créer des compartiments et comment transférer des données vers Amazon S3, consultez le Guide du l'utilisateur Amazon Simple Storage Service.
2. Créez un dossier sur votre cluster Hadoop pour vos fichiers source en entrant une commande similaire à celle qui suit :

   ```
   mkdir SourceFiles
   ```

3. Copiez vos fichiers source à partir d'Amazon S3 vers le nœud principal en tapant une commande similaire à celle qui suit :

   ```
   hadoop fs -get s3://amzn-s3-demo-bucket/SourceFiles SourceFiles
   ```

Pour créer des fichiers binaires avec les optimisations nécessaires

La façon dont vous créez vos fichiers binaires dépend de plusieurs facteurs. Suivez les instructions correspondant aux outils d'installation que vous utilisez pour installer et configurer votre environnement. Vous pouvez utiliser les commandes de spécification du système Hadoop pour obtenir les informations relatives au cluster, afin de déterminer comment installer votre environnement.

Pour identifier les spécifications du système

- Utilisez les commandes suivantes pour vérifier l'architecture que vous utilisez pour créer vos fichiers binaires.

  a. Pour afficher la version de Debian, saisissez la commande suivante :

  ```
  master$ cat /etc/issue
  ```

  La sortie ressemble à ce qui suit.

  ```
  Debian GNU/Linux 5.0
  ```

  b. Pour afficher le nom DNS public et la taille du processeur, saisissez la commande suivante :

  ```
  master$ uname -a
  ```

  La sortie ressemble à ce qui suit.

  ```
  Linux domU-12-31-39-17-29-39.compute-1.internal 2.6.21.7-2.fc8xen #1 SMP Fri
   Feb 15 12:34:28 EST 2008 x86_64 GNU/Linux
  ```

  c. Pour afficher la vitesse du processeur, saisissez la commande suivante :

  ```
  master$ cat /proc/cpuinfo
  ```

  La sortie ressemble à ce qui suit.

  ```
  processor : 0
  vendor_id : GenuineIntel
  model name : Intel(R) Xeon(R) CPU E5430 @ 2.66GHz
  flags : fpu tsc msr pae mce cx8 apic mca cmov pat pse36 clflush dts acpi mmx
   fxsr sse sse2 ss ht tm syscall nx lm constant_tsc pni monitor ds_cpl vmx est
   tm2 ssse3 cx16 xtpr cda lahf_lm
  ```

```
...
```

Une fois que vos fichiers binaires sont créés, vous pouvez copier les fichiers dans Amazon S3.

Pour copier des fichiers binaires à partir du nœud principal dans Amazon S3

- Entrez la commande suivante pour copier les fichiers binaires dans votre compartiment Amazon S3 :

```
hadoop fs -put BinaryFiles s3://amzn-s3-demo-bucket/BinaryDestination
```

Traitement des données avec Streaming

Le streaming Hadoop est un utilitaire qui est fourni avec Hadoop pour vous permettre de développer des MapReduce fichiers exécutables dans d'autres langages que Java. Streaming est mis en œuvre sous la forme d'un fichier JAR, afin que vous puissiez l'exécuter à partir de la ligne de commande ou de l'API Amazon EMR, tout comme un fichier JAR standard.

Cette section décrit comment utiliser le streaming avec Amazon EMR.

> **Note**
>
> Le streaming Apache Hadoop est un outil indépendant. En tant que tel, ses fonctions et ses paramètres ne sont pas tous décrits ici. Pour plus d'informations sur le streaming Hadoop, accédez à http://hadoop.apache. org/docs/stable/hadoop-streaming/HadoopStreaming.html.

Utilisation de l'utilitaire de streaming Hadoop

Cette section décrit comment utiliser l'utilitaire de streaming Hadoop.

Processus Hadoop

1	Ecrivez votre exécutable mappeur et réducteur dans le langage de programmation de votre choix. Suivez les instructions dans la documentation Hadoop pour écrire vos exécutable de diffusion en continu. Les programmes doivent lire leur entrée à partir de données d'entrée et de sortie standard par le biais de sortie standard. Par défaut, chaque ligne d'entrée/

	sortie représente un enregistrement et le premier onglet sur chaque ligne est utilisé comme séparateur entre la clé et la valeur.
2	Testez vos exécutables localement et chargez-les sur Amazon S3.
3	Utilisez l'interface de ligne de commande d'Amazon EMR ou la console d'Amazon EMR pour exécuter votre application.

Chaque script mappeur se lance comme un processus distinct dans le cluster. Chaque exécutable de réducteur remplace le résultat de l'exécutable du mappeur par la sortie de données du flux de travail.

Les paramètres `input`, `output`, `mapper`, et `reducer` sont requis par la plupart des applications de streaming. Le tableau suivant décrit ces paramètres et d'autres, facultatifs.

Paramètre	Description	Obligatoire
-input	Emplacement des données d'entrée sur Amazon S3. Type : String Par défaut : aucun Contrainte : URI. Si aucun protocole n'est spécifié, il utilise alors le système de fichiers par défaut du cluster.	Oui
-output	Emplacement sur Amazon S3 où Amazon EMR charge les données traitées. Type : String Par défaut : aucun Contrainte : URI Par défaut : si aucun emplacement n'est spécifié, Amazon EMR charge les données à l'emplacement spécifié par `input`.	Oui
-mapper	Nom de l'exécutable du mappeur.	Oui

Paramètre	Description	Obligatoire
	Type : String Par défaut : aucun	
-reducer	Nom de l'exécutable du réducteur. Type : String Par défaut : aucun	Oui
-cacheFile	Un emplacement Amazon S3 qui contient des fichiers que Hadoop peut copier dans votre répertoire de travail local (principalement pour améliorer les performances). Type : String Par défaut : aucun Contraintes : [URI]#[nom symlink à créer dans répertoire de travail]	Non
-cacheArchive	Fichier JAR à extraire dans le répertoire de travail Type : String Par défaut : aucun Contraintes : [URI]#[nom de répertoire symlink à créer dans répertoire de travail	Non
-combiner	Combine les résultats Type : String Par défaut : aucun Contraintes : nom de classe Java	Non

L'exemple de code suivant est un exécutable de mappeur écrit dans Python. Ce script fait partie de l' WordCount exemple d'application.

```
#!/usr/bin/python
import sys

def main(argv):
  line = sys.stdin.readline()
  try:
    while line:
      line = line.rstrip()
      words = line.split()
      for word in words:
        print "LongValueSum:" + word + "\t" + "1"
      line = sys.stdin.readline()
  except "end of file":
    return None
if __name__ == "__main__":
  main(sys.argv)
```

Soumission d'une étape de streaming

Cette section présente les concepts de base de la soumission d'une étape de streaming à un cluster. Une application de streaming lit l'entrée depuis l'entrée standard, puis exécute un script ou un exécutable (appelé un mappeur) sur chaque entrée. Le résultat de chacune des entrées est enregistré localement, généralement sur une partition de système de fichiers distribué Hadoop (HDFS). Une fois que toutes les entrées sont traitées par le mappeur, un second script ou exécutable (appelé un réducteur) traite les résultats du mappeur. Les résultats du réducteur sont envoyés vers la sortie standard. Vous pouvez enchaîner une série d'étapes de streaming, où le résultat d'une étape devient l'entrée d'une autre.

Le mappeur et le réducteur peuvent être chacun référencés en tant que fichier ou vous pouvez fournir une classe Java. Vous pouvez implémenter le mappeur et le réducteur dans n'importe lequel des langages pris en charge, y compris Ruby, Perl, Python, PHP ou Bash.

Soumission d'une étape de streaming à l'aide de la console

Cet exemple décrit comment utiliser la console Amazon EMR pour soumettre une étape streaming à un cluster en cours d'exécution.

Pour soumettre une étape de streaming

1. Ouvrez la console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com
2. Dans la Liste de clusters, sélectionnez le nom de votre cluster.
3. Faites défiler l'affichage jusqu'à la section Étapes et développez-la, puis choisissez Ajouter une étape.
4. Dans la boîte de dialogue Ajouter une étape :
   - Pour Type d'étape, choisissez Programme de streaming.
   - Pour Nom, acceptez le nom par défaut (Programme de streaming) ou saisissez un nouveau nom.
   - Pour Mappeur, saisissez l'emplacement de votre classe de mappeur, ou accédez à ce dernier via le navigateur, dans Hadoop ou dans un compartiment S3 dans lequel l'exécutable de mappeur, tel qu'un programme Python, réside. La valeur du chemin doit être au format*BucketName*/*path*/*MapperExecutable*.
   - Pour Reducer (Réducteur), saisissez l'emplacement de votre classe de mappeur, ou accédez à ce dernier via le navigateur, dans Hadoop, ou dans un compartiment S3 dans lequel l'exécutable du réducteur, tel qu'un programme Python, réside. La valeur du chemin doit être au format*BucketName*/*path*/*MapperExecutable*. Amazon EMR prend en charge le mot clé spécial aggregate. Pour plus d'informations, consultez la bibliothèque Aggregate fournie par Hadoop.
   - Pour Emplacement S3 d'entrée, saisissez l'emplacement de vos données d'entrée ou accédez à ce dernier via le navigateur.
   - Pour Emplacement S3 de sortie, saisissez le nom du compartiment de sortie Amazon S3 ou naviguez à ce dernier.
   - Pour Arguments, laissez le champ vide.
   - Pour Action sur échec, acceptez l'option par défaut (Continuer).
5. Choisissez Ajouter. L'étape s'affiche dans la console avec le statut En suspens.
6. Le statut de l'étape passe de Pending (En suspens) à Running (En cours d'exécution) puis à Completed (Terminé), au fur et à mesure de son exécution. Pour mettre à jour l'état, choisissez l'icône d'actualisation au-dessus de la colonne Actions.

AWS CLI

Ces exemples montrent comment utiliser l'étape AWS CLI pour créer un cluster et soumettre une étape de streaming.

Pour créer un cluster et soumettre une étape de streaming à l'aide de l' AWS CLI

- Pour créer un cluster et soumettre une étape de streaming à l'aide de l' AWS CLI, tapez la commande suivante et *myKey* remplacez-la par le nom de votre paire de EC2 clés. Notez que votre argument pour `--files` doit être le chemin Amazon S3 vers l'emplacement de votre script, et les arguments pour `-mapper` et `-reducer` doivent être les noms des fichiers de script respectifs.

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 --
applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=STREAMING,Name="Streaming Program",ActionOnFailure=CONTINUE,Args=[--
files,pathtoscripts,-mapper,mapperscript,-reducer,reducerscript,aggregate,-
input,pathtoinputdata,-output,pathtooutputbucket]
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

Lorsque vous spécifiez le nombre d'instances sans utiliser le paramètre `--instance-groups`, un seul nœud maître est lancé et les instances restantes sont lancées en tant que nœuds principaux. Tous les nœuds utilisent le type d'instance spécifié dans la commande.

> **Note**
>
> Si vous n'avez pas encore créé le rôle de service Amazon EMR par défaut et le profil d' EC2 instance Amazon EMR, tapez aws `emr create-default-roles` pour les créer avant de taper la `create-cluster` sous-commande.

Pour plus d'informations sur l'utilisation des commandes Amazon EMR dans le AWS CLI, consultez. https://docs.aws.amazon.com/cli/latest/reference/emr

Traitement de données à l'aide d'un fichier JAR personnalisé

Un fichier JAR personnalisé exécute un programme Java compilé que vous pouvez charger sur Amazon S3. Vous devez compiler le programme en fonction de la version d'Hadoop que vous voulez lancer, et soumettre une étape `CUSTOM_JAR` à votre cluster Amazon EMR. Pour plus d'informations sur la compilation d'un fichier JAR, consultez Créer des fichiers binaires à l'aide d'Amazon EMR.

Pour plus d'informations sur la création d'une MapReduce application Hadoop, consultez le MapReduce Tutoriel dans la documentation Apache Hadoop.

Rubriques

- Soumission d'une étape JAR personnalisée

Soumission d'une étape JAR personnalisée

Un fichier JAR personnalisé exécute un programme Java compilé que vous pouvez charger sur Amazon S3. Vous devez compiler le programme en fonction de la version d'Hadoop que vous voulez lancer, et soumettre une étape `CUSTOM_JAR` à votre cluster Amazon EMR. Pour plus d'informations sur la compilation d'un fichier JAR, consultez Créer des fichiers binaires à l'aide d'Amazon EMR.

Pour plus d'informations sur la création d'une MapReduce application Hadoop, consultez le MapReduce Tutoriel dans la documentation Apache Hadoop.

Cette section présente les principes de base de la soumission d'une étape JAR personnalisée dans Amazon EMR. La soumission d'une étape JAR personnalisée vous permet d'écrire un script pour traiter vos données à l'aide du langage de programmation Java.

Soumettre une étape JAR personnalisée à l'aide de la console

Cet exemple décrit comment utiliser la console Amazon EMR pour soumettre une étape JAR personnalisée à un cluster en cours d'exécution.

Pour soumettre une étape JAR personnalisée à l'aide de la console

1. Ouvrez la console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com

2. Dans la Liste de clusters, sélectionnez le nom de votre cluster.
3. Faites défiler l'affichage jusqu'à la section Étapes et développez-la, puis choisissez Ajouter une étape.
4. Dans la boîte de dialogue Ajouter une étape :
   - Pour Type d'étape, choisissez JAR personnalisé.
   - Pour Nom, acceptez le nom par défaut (JAR personnalisé) ou tapez un nouveau nom.
   - Pour JAR S3 location (Emplacement S3 JAR), tapez ou accédez à l'emplacement de votre fichier JAR. L'emplacement JAR peut-être un chemin dans S3 ou une classe Java entièrement qualifiée dans le classpath.
   - Pour Arguments, tapez les arguments requis sous forme de chaînes séparées par des espaces ou laissez le champ vide.
   - Pour Action sur échec, acceptez l'option par défaut (Continuer).
5. Choisissez Ajouter. L'étape s'affiche dans la console avec le statut En suspens.
6. Le statut de l'étape passe de Pending (En suspens) à Running (En cours d'exécution) puis à Completed (Terminé), au fur et à mesure de son exécution. Pour mettre à jour l'état, choisissez l'icône d'actualisation au-dessus de la colonne Actions.

Lancement d'un cluster et soumission d'une étape JAR personnalisée à l'aide de l' AWS CLI

Pour lancer un cluster et soumettre une étape JAR personnalisée à l'aide de l' AWS CLI

Pour lancer un cluster et soumettre une étape JAR personnalisée à l'aide de l' AWS CLI, tapez la `create-cluster` sous-commande avec le `--steps` paramètre.

- Pour lancer un cluster et soumettre une étape JAR personnalisée, tapez la commande suivante, remplacez-la *myKey* par le nom de votre paire de EC2 clés et remplacez *amzn-s3-demo-bucket* par le nom de votre bucket.

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=CUSTOM_JAR,Name="Custom JAR
 Step",ActionOnFailure=CONTINUE,Jar=pathtojarfile,Args=["pathtoinputdata","pathtooutputbuck
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

Lorsque vous spécifiez le nombre d'instances sans le paramètre `--instance-groups`, un seul nœud primaire est lancé et les instances restantes sont lancées en tant que nœuds principaux. Tous les nœuds utilisent le type d'instance que vous spécifiez dans la commande.

> **Note**
>
> Si vous n'avez pas encore créé le rôle de service et le profil d'instance par défaut et le profil d'EC2 instance, tapez `aws emr create-default-roles` pour les créer avant de taper la `create-cluster` sous-commande.

Pour plus d'informations sur l'utilisation des commandes Amazon EMR dans le AWS CLI, consultez. https://docs.aws.amazon.com/cli/latest/reference/emr

Dépendances tierces

Parfois, il peut être nécessaire de l'inclure dans le MapReduce chemin de classe JARs à utiliser avec votre programme. Vous avez deux options pour cela :

- Incluez `--libjars s3://`*`URI_to_JAR`* dans les options de l'étape pour la procédure dans Lancement d'un cluster et soumission d'une étape JAR personnalisée à l'aide de l' AWS CLI.
- Lancez le cluster avec un paramètre `mapreduce.application.classpath` modifié dans `mapred-site.xml`. Utilisez la classification de configuration `mapred-site`. Pour créer le cluster avec l'étape à l'aide de l' AWS CLI, cela se présente comme suit :

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--instance-type m5.xlarge --instance-count 2  --ec2-attributes KeyName=myKey \
```

```
--steps Type=CUSTOM_JAR,Name="Custom JAR
 Step",ActionOnFailure=CONTINUE,Jar=pathtojarfile,Args=["pathtoinputdata","pathtooutputbucket
 \
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json
```

`myConfig.json`:

```
[
    {
      "Classification": "mapred-site",
      "Properties": {
        "mapreduce.application.classpath": "path1,path2"
      }
    }
  ]
```

La liste séparée par des virgules des chemins d'accès doit être ajoutée au chemin de classe pour la JVM de chaque tâche.

Lire les objets restaurés

Avec les versions 7.2.0 et supérieures d'Amazon EMR, vous pouvez lire les objets Glacier restaurés depuis l'emplacement S3 de la table contenant le protocole. `S3A` Les moteurs des versions précédentes ne font pas de distinction entre les fichiers Glacier et Glacier Deep Archive, ce qui signifie que vous en obteniez un `AmazonS3Exception` si vous essayiez d'accéder à un fichier Glacier en cours avec`S3A`. Cette opération de lecture ignore les fichiers Glacier archivés s'ils sont toujours en cours de restauration. Pour activer ce comportement, utilisez le paramètre`fs.s3a.glacier.read.restored.objects`. Ce paramètre peut prendre trois valeurs :

- READ_ALL : cette valeur indique qu'Amazon EMR ne doit pas prendre en compte les classes de stockage extraites d'Amazon S3. Il s'agit du comportement de par défaut.
- SKIP_ALL_GLACIER : cette valeur indique qu'Amazon EMR doit ignorer tous les objets S3 marqués avec la classe de stockage Glacier et récupérer tous les autres objets. Il s'agit du comportement par défaut d'Amazon Athena en ce qui concerne les objets Glacier.
- READ_RESTORED_GLACIER_OBJECTS : cette valeur indique qu'Amazon EMR doit vérifier l'état restauré de l'objet Glacier. Si Amazon EMR peut restaurer l'objet, vous pouvez le lire comme un objet S3 normal. Dans le cas contraire, Amazon EMR ignore l'objet de S3 Glacier

Exemples

Spark

Pour lire les objets restaurés lorsque vous utilisez Spark, utilisez la configuration suivante :

```
--conf spark.hadoop.fs.s3a.glacier.read.restored.objects=<value>
```

Si vous utilisez spark-sql, utilisez plutôt la configuration suivante ;

```
spark-sql --conf spark.hadoop.fs.s3a.glacier.read.restored.objects=<value>
```

Flink

Si vous utilisez Flink, vous pouvez définir la configuration dans le fichier `flink-conf.yaml`

```
fs.s3a.glacier.read.restored.objects: <value>
```

Vous pouvez également définir le `flink-conf` classement :

```
[
 {
  "Classification": "flink-conf",
  "Properties": {
  "fs.s3a.glacier.read.restored.objects":"<value>"
  }
 }
]
```

Hive

Si vous utilisez Hive, définissez la configuration dans le `hive-site.xml` fichier.

```
<property>
 <name>fs.s3a.glacier.read.restored.objects</name>
 <value><value></value>
</property>
```

Vous pouvez également utiliser la CLI Hive pour définir la propriété `--hiveconf` :

```
hive --hiveconf fs.s3a.glacier.read.restored.objects=<value>
```

Considérations

Lorsque vous lisez des objets restaurés depuis S3 Glacier, tenez compte des considérations suivantes :

- Vous ne pouvez lire les objets restaurés que si vous utilisez le `S3A` schéma ou le `S3AFileSystem` pour accéder aux données.
- Lorsque vous lisez un objet Glacier restauré, Amazon EMR ne le restaure pas lui-même. Pour ce faire, vous devez utiliser le AWS CLI ou le AWS SDK.

Activer la détection non uniforme des accès à la mémoire pour les conteneurs YARN

Avec les versions 6.x et ultérieures d'Amazon EMR, vous pouvez utiliser l'accès mémoire non uniforme (NUMA) pour le multitraitement de vos données sur des clusters. NUMA est un modèle de conception de mémoire informatique dans lequel le processeur peut accéder à sa propre mémoire locale plus rapidement que la mémoire d'un autre processeur ou partagée entre processeurs. Les conteneurs YARN offrent de meilleures performances avec NUMA car ils peuvent se lier à un nœud NUMA spécifique qui gère toutes les allocations de mémoire ultérieures. Cela réduit le nombre de fois que votre cluster doit accéder à la mémoire distante.

Vous pouvez activer le support NUMA pour le conteneur YARN lorsque la machine du composant master est un nœud multi-NUMA. Pour vérifier si un composant master est un nœud NUMA unique ou multi-NUMA, exécutez la commande suivante.

```
lscpu | grep -i numa
NUMA node(s): 2
```

En général, les instances dont la taille est supérieure à 12 fois possèdent deux nœuds NUMA. Cela ne s'applique pas aux instances en métal.

Pour activer la prise en compte de NUMA pour les conteneurs YARN

1. Utilisez la configuration `yarn-site` suivante dans votre cluster Amazon EMR 6.x.

```
[
{
 "classification":"yarn-site",
```

```
   "properties":{
    "yarn.nodemanager.linux-container-executor.nonsecure-mode.local-user":"yarn",
    "yarn.nodemanager.linux-container-executor.group":"yarn",
    "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor",
    "yarn.nodemanager.numa-awareness.enabled":"true",
    "yarn.nodemanager.numa-awareness.numactl.cmd":"/usr/bin/numactl",
    "yarn.nodemanager.numa-awareness.read-topology":"true"
   },
  "configurations":[]
 }
 ]
```

2. Fournissez l'action d'amorçage suivante dans votre cluster.

```
#!/bin/bash

sudo yum -y install numactl
echo 1 | sudo tee /proc/sys/kernel/numa_balancing

echo "banned.users=mapred,bin,hdfs" >> /etc/hadoop/conf/container-executor.cfg
rm -rf /var/log/hadoop-yarn/
sudo chown -R yarn:hadoop /var/log/hadoop-yarn/
sudo chmod 755 -R /var/log/hadoop-yarn/

sudo chmod 6050 /etc/hadoop/conf/container-executor.cfg

mkdir /mnt/yarn && sudo chmod 755 -R /mnt/yarn && sudo chown -R yarn:hadoop /mnt/
yarn
mkdir /mnt1/yarn && sudo chmod 755 -R /mnt1/yarn && sudo chown -R yarn:hadoop /
mnt1/yarn
mkdir /mnt2/yarn && sudo chmod 755 -R /mnt2/yarn && sudo chown -R yarn:hadoop /
mnt2/yarn
```

3. Chaque conteneur doit être conscient de l'existence de NUMA. Vous pouvez informer la Java Virtual machine (JVM) de chaque conteneur par un drapeau NUMA. Par exemple, pour demander à la JVM d'utiliser NUMA dans une tâche MapReduce, ajoutez les propriétés suivantes dans `mapred-site.xml`.

```
<property>
 <name>mapreduce.reduce.java.opts</name>
 <value>-XX:+UseNUMA</value>
</property>
```

```
<property>
 <name>mapreduce.map.java.opts</name>
 <value>-XX:+UseNUMA</value>
</property>
```

4. Pour vérifier que vous avez activé NUMA, effectuez une recherche dans l'un des fichiers journaux de NodeManager à l'aide de la commande suivante.

   ```
   grep "NUMA resources allocation is enabled," *
   ```

   Pour vérifier que des ressources de NUMA nœud NodeManager ont été attribuées à un conteneur, effectuez une recherche dans le NodeManager journal à l'aide de la commande suivante, en *<container_id>* remplaçant par votre propre ID de conteneur.

   ```
   grep "NUMA node" | grep <container_id>
   ```

Emballage en poubelle YARN

À partir de la version 7.9.0 d'Amazon EMR, la politique d'emballage des conteneurs est désormais disponible pour le planificateur de capacité YARN, qui repose sur la politique de placement multi-nœuds de YARN. Bien que la fonctionnalité soit désactivée par défaut, lorsqu'elle est activée, YARN donne la priorité au remplissage d'un seul nœud avec des conteneurs avant de l'étendre à d'autres nœuds du cluster, tout en respectant un seuil d'emballage prédéfini défini par la configuration. `yarn.scheduler.capacity.multi-node-placement.container.bin-packing.percentage`

La politique d'emballage des conteneurs offre plusieurs avantages par rapport à la stratégie uniforme d'allocation des conteneurs par défaut :

- Il réduit la fragmentation des ressources du cluster.
- Il peut accélérer les opérations de réduction des clusters en lançant des conteneurs sur un nombre limité de nœuds lorsque des ressources sont disponibles sur ces nœuds, laissant ainsi les autres nœuds inactifs, qui peuvent ensuite être réduits, ce qui permet de réaliser de meilleures économies lors de la mise à l'échelle dynamique d'un cluster.

Activez la fonctionnalité

Pour activer la fonctionnalité d'emballage des conteneurs dans Amazon EMR, vous pouvez ajouter la classification de site YARN suivante :

```
[
 {
  "Classification": "yarn-site",
  "Properties": {
  "yarn.scheduler.capacity.multi-node-placement.container.bin-packing.percentage":
 "integer value from 1-100"
  }
 }
]
```

Considérations

- Cette fonctionnalité est exclusivement disponible pour le planificateur de capacité YARN.
- L'activation de cette fonctionnalité active automatiquement la stratégie de planification du placement multi-nœuds YARN.
- Il peut y avoir une dégradation potentielle des performances en raison de l'utilisation concentrée des ressources sur un nombre limité de nœuds.
- Grâce à cette fonctionnalité, les politiques d'auto-scaling personnalisées permettent de meilleures opérations de mise à l'échelle par rapport aux politiques de dimensionnement gérées.

Historique des versions de Hadoop

Le tableau suivant répertorie la version de Hadoop incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version de Hadoop

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-7.9.0	3.4.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7,8.0	3.4.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7,7.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resour

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
		cemanager, hadoop-yarn-timeline-server
emr-7.6.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.5.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-7.4.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.3.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-7.2.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.36.2	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-7.1.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.0.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.15.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.14.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.13.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.12.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.11.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.11.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.10.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.10.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.9.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.9.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.8.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.8.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.7.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.36.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.36.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.6.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.35.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.5.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.4.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.3.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.3.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.2.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.2.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.1.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.1.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.0.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-6.0.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.34.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.33.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.33.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.32.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.32.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.31.1	2.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.31.0	2.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.30.2	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.30.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.30.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
mr-5.29.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.28.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.28.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.27.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.27.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.26.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.25.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.24.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.24.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.23.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.23.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.22.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.21.2	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.21.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.21.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.20.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.20.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.19.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.19.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.18.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.18.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.17.2	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.17.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.17.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.16.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.16.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.15.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.15.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.14.2	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.14.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.14.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.13.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.13.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.12.3	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.12.2	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.12.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.12.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.11.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.11.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.11.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.10.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.10.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.9.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.9.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.8.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.8.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.8.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.8.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.7.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.7.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.6.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.6.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.5.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.5.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.5.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.5.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.5.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.4.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.4.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.3.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.3.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.3.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.2.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.2.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.2.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.2.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.1.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.1.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.0.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-5.0.2	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.0.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.0.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-4.9.6	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.5	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-4.9.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-4.8.5	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-4.8.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-4.7.4	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.3	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.2	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-4.7.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.6.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-4.6.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.5.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.4.0	2.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-4.3.0	2.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.2.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.1.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Étiquette de version Amazon EMR	Version de Hadoop	Composants installés avec Hadoop
emr-4.0.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Notes de mise à jour de Hadoop par version

Rubriques

- Amazon EMR 7.9.0 - Notes de mise à jour de Hadoop
- Amazon EMR 7.8.0 - Notes de mise à jour de Hadoop
- Amazon EMR 7.7.0 - Notes de mise à jour de Hadoop
- Amazon EMR 7.6.0 - Notes de mise à jour de Hadoop
- Amazon EMR 7.5.0 - Notes de mise à jour de Hadoop
- Amazon EMR 7.4.0 - Notes de mise à jour de Hadoop
- Amazon EMR 7.3.0 - Notes de mise à jour de Hadoop
- Amazon EMR 7.2.0 - Notes de mise à jour de Hadoop
- Amazon EMR 6.6.0 – Notes de mise à jour de Hadoop

Amazon EMR 7.9.0 - Notes de mise à jour de Hadoop

Amazon EMR 7.9.0 - Modifications apportées à Hadoop

Type	Description
Nouvelle fonction	Cartographie automatique de la configuration : la configuration du système de fichiers

Type	Description
	EMRFS est automatiquement appliquée à la configuration du système de fichiers S3A correspondante, ce qui permet une migration fluide d'EMRFS vers S3A.
Nouvelle fonction	Emballage de conteneurs YARN : politique de planification pour une réduction d'échelle agressive
Rétroportage	YARN-11752 : Planificateur global : améliorer le temps d'allocation des conteneurs
Correctif de bogue.	S3A : Corriger le IAMCredentials fournisseur renvoyant des informations d'identification expirées
Correctif de bogue.	Fixer `yarn.log.server.url` la valeur de configuration pour les clusters avec le chiffrement en transit activé

Amazon EMR 7.8.0 - Notes de mise à jour de Hadoop

Amazon EMR 7.8.0 - Modifications apportées à Hadoop

Type	Description
Upgrade	Hadoop est mis à niveau vers la version 3.4.1.

Amazon EMR 7.7.0 - Notes de mise à jour de Hadoop

Amazon EMR 7.7.0 - Modifications apportées à Hadoop

Type	Description
Nouvelle fonction	Optimisez l' GlobStatus appel S3A avec la liste des préfixes S3
Rétroportage	YARN-7327 : Activer la planification asynchrone par défaut pour le planificateur de capacité
Rétroportage	YARN-10058 : Gérez les exceptions non détectées pour les threads de planification asynchrone afin d'éviter le blocage du planificateur
Rétroportage	YARN-11732 : Corrigez un NPE potentiel lorsque vous appelez #reservedContainer pour SchedulerNode CapacityScheduler
Rétroportage	YARN-11560 : Correction d'un bogue NPE lorsque le multi-nœuds est activé avec une planification asynchrone
Rétroportage	YARN-11191 : Correction d'un blocage potentiel dans RefreshQueues GlobalScheduler
Rétroportage	YARN-11041 : Remplacez toutes les occurrences de QueuePath par la nouvelle classe QueuePath
Rétroportage	

Type	Description
	YARN-11660 : Corriger la régression des performances pour SingleConstraintAp pPlacementAllocator
Rétroportage	HADOOP-19116 : Mise à jour du client zookeeper 3.8.4 en raison du CVE-2024-23944.
Rétroportage	HADOOP-19115 : mise à niveau vers la version nimbus-jose-jwt 9.37.2 en raison de CVE-2023-52428.
Rétroportage	HADOOP-19024 : Utiliser bouncycastle jdk18 1.77
Rétroportage	HADOOP-19123 : Mise à jour de commons-configuration2 2.10.1 à cause de CVE
Rétroportage	HADOOP-19114 : Mise à niveau vers commons-compress 1.26.1 en raison de CVEs
Rétroportage	HADOOP-19237 : Mise à niveau vers dnsjava 3.6.1 en raison de CVEs
Nouvelle fonction	Ajouter l'audit des demandes S3 à S3A
Rétroportage	HADOOP-17609 : Rendre le support SM4 optionnel pour le code natif d'OpenSSL
Rétroportage	HADOOP-18583 : hadoop checknative ne parvient pas à charger openssl 3.x

Type	Description
Nouvelle fonction	Ajout de la prise en charge des mappages de rôles S3A

Amazon EMR 7.7.0 - Fonctionnalités de Hadoop

- La planification asynchrone des conteneurs est la stratégie de planification par défaut du planificateur de capacité, conçue pour optimiser la vitesse d'allocation des conteneurs.
- Le système de fichiers S3A introduit une optimisation pour les appels d'état globaux en utilisant la liste des préfixes S3 pour accélérer les opérations de liste. Par défaut, cette fonctionnalité est désactivée et peut être activée `fs.s3a.prefix.listing.in.glob.status.enabled=true` en configurant dans le fichier core-site.xml. Lorsqu'elle est activée, l'optimisation permet le filtrage côté serveur pour les appels globstatus`fs.globstatus("s3://bucket/a*")`, par exemple en améliorant les performances des listes en répertoriant uniquement les objets commençant par. `"a"`
- Ajoutez l'audit des demandes S3 à S3A. Lorsque cette option est activée, les informations provenant de fileSystemOwner l'objet sont utilisées pour renseigner la chaîne UserAgent avec les champs utilisateur et groupe d'utilisateurs effectuant les demandes S3.
- S3A ajoute la prise en charge des mappages de rôles, ce qui permet de déterminer le rôle IAM à utiliser en fonction des utilisateurs, des groupes ou des préfixes S3.

Amazon EMR 7.6.0 - Notes de mise à jour de Hadoop

Amazon EMR 7.6.0 - Modifications apportées à Hadoop

Type	Description
Nouvelle fonction	Cadre de la vente d'informations d'identification S3A
Nouvelle fonction	Support du S3A Magic Committer pour Hive
Rétroportage	

Type	Description
	HADOOP-19309 : S3A : l' CopyFromLocalFile opération échoue lorsque le fichier source ne contient pas de schéma de fichier
Rétroportage	HADOOP-19050 S3A:Supporte les autorisations d'accès S3
Rétroportage	HADOOP-18708 S3A : Support du chiffrement côté client S3 (CSE)

Amazon EMR 7.6.0 - Fonctionnalités de Hadoop

- Les paramètres de configuration par défaut pour `yarn.scheduler.increment-allocation-mb` et `yarn.scheduler.minimum-allocation-mb` ont été modifiés de 32 à 1, ce qui désactive efficacement la normalisation des ressources des conteneurs dans YARN.
- Le système de fichiers S3A de Hadoop a été amélioré pour prendre en charge la vente d'informations d'identification au niveau des demandes. Grâce à cette fonctionnalité, chaque demande Amazon S3 effectuée utilisera un`AwsCredentialsProvider`. Cela garantit une flexibilité et une sécurité améliorées en fournissant des informations d'identification dynamiques pour chaque demande.
- Le système de fichiers S3A de Hadoop prend en charge le chiffrement S3 côté client.

Amazon EMR 7.5.0 - Notes de mise à jour de Hadoop

Amazon EMR 7.5.0 - Modifications apportées à Hadoop

Type	Description
Correctif de bogue.	A commenté fs.file.impl sur une valeur vide.
Rétroportage	

Type	Description
	HADOOP-19286 : Support de l'accès interrégional au S3A lorsque S3 est défini region/endpoint
Amélioration	Paramètre de configuration automatique de la région S3 pour le connecteur S3A sur EMR-EC2
Amélioration	Réduisez le nombre d' HeadObject appels dans le S3A

Avec la sortie d'Amazon EMR 7.5, le connecteur S3A de Spark affiche des performances de lecture comparables à celles d'EMRFS, comme en témoignent les tests de performance basés sur un ensemble de données de 3 To pour parquet TPC-DS.

Amazon EMR 7.5.0 - Fonctionnalités de Hadoop

- `fs.s3a.endpoint.region`La configuration de la région S3 est automatiquement définie sur la région dans laquelle le cluster EMR est lancé avec le connecteur S3A pour le déploiement de l'EMR. EC2
- L'accès aux régions entre compartiments Amazon S3 est activé par défaut pour le connecteur S3A. Il peut être modifié par réglage`fs.s3a.cross.region.access.enabled=`*`true or false`*.

Amazon EMR 7.4.0 - Notes de mise à jour de Hadoop

Amazon EMR 7.4.0 - Modifications apportées à Hadoop

Type	Description
Upgrade	La version Hadoop est mise à niveau vers la version 3.4.0, reportez-vous aux notes de mise à jour de l'OSS.

Type	Description
Correctif de bogue.	Corriger les métriques négatives relatives au fil en attente et au fil alloué pour FairScheduler
Correctif de bogue.	YARN-11702 : Fixez le fil sur les conteneurs d'allocation
Correctif de bogue.	Améliorez la gestion des conditions de course lors de la réduction de la taille des nœuds
Amélioration	HADOOP-18679 : Ajout d'une API pour la suppression de fichiers bulk/paged
Amélioration	HADOOP-19203 : L'API BulkDelete Wrappedio doit être définie comme étant non contrôlée IOEs IOExceptions
Amélioration	HADOOP-19205 : S3A : initialization/close plus lent qu'avec le SDK v1
Amélioration	HADOOP-19161 : S3A : option fs.s3a.performance.flags pour récupérer la liste des indicateurs de performance
Amélioration	HADOOP-19072 : S3A : étendez les optimisations sur les stores avec fs.s3a.performance.flags pour mkdir

Amazon EMR 7.4.0 - Fonctionnalités de Hadoop

Consultez la liste suivante pour découvrir les nouvelles fonctionnalités Hadoop dans Amazon EMR 7.4.0.

- Les valeurs de configuration par défaut ont été affinées pour des performances optimales :
  - mapreduce.input.fileinputformat.list-status.num-threads=10 — C'est une hausse par rapport à 1.
  - fs.s3a.block.size=64M — Cela représente une augmentation par rapport à 32M.
  - fs.s3a.multipart.size=128M — Cela représente une hausse par rapport à 64M.
- Out-of-the-box optimisations améliorant les performances pour accélérer les MapReduce tâches avec le système de fichiers S3A.

Amazon EMR 7.3.0 - Notes de mise à jour de Hadoop

Amazon EMR 7.3.0 - Modifications apportées à Hadoop

Type	Description
Nouvelle fonction	HADOOP-19197 : Support S3A : contexte de chiffrement AWS KMS
Nouvelle fonction	HADOOP-18980 : Remappage du fournisseur d'informations d'identification S3A : rendre extensible
Amélioration	HADOOP-18808 : LogExactlyOnce pour ajouter une méthode debug ()

Amazon EMR 7.3.0 - Fonctionnalités de Hadoop

Consultez la liste suivante pour découvrir les nouvelles fonctionnalités Hadoop dans Amazon EMR 7.3.0.

- Par défaut, les clusters dont le chiffrement en transit est activé via la configuration de sécurité exécuteront le protocole TLS 1.3 pour Hadoop YARN (Resource Manager, Node Manager, Timeline Server et WebApplicationProxy), HDFS (, NameNode DataNode, Journal Node et DFSZKFailover Controller) et Map Reduce (ports JobHistory Server et Shuffle). MapReduce

Amazon EMR 7.2.0 - Notes de mise à jour de Hadoop

Amazon EMR 7.2.0 - Modifications apportées à Hadoop

Type	Description
Nouvelle fonction	HADOOP-18850 : S3A : activer le chiffrement double couche côté serveur à l'aide de clés AWS KMS
Nouvelle fonction	Support S3 Glacier pour lire les objets restaurés
Amélioration	Dimensionnement dynamique du nombre maximal de connexions S3A en fonction des cœurs lors de l'exécution de Spark Jobs
Amélioration	HADOOP-18797 : Support des écritures simultanées avec le S3A Magic Committer
Amélioration	HADOOP-19047 : InMemory Support et suivi des S3A Magic Commits
Amélioration	Ajouter les méthodes « Export », « Publish » et « » AbortPublish dans MagicV2 Committer
Amélioration	Implémenter ReadFullyIntoBuffers FileSystem l'API dans S3A
Correctif de bogue.	HADOOP-18793 : S3A StagingCommitter ne nettoie pas le répertoire staging-uploads
Upgrade	

Type	Description
	HADOOP-19008 : S3A : mise à jour de la version aws-sdk vers la version 2.21.41
Upgrade	HADOOP-18613 : Mise à niveau vers la version 3.8.3 ZooKeeper

Amazon EMR 7.2.0 - Fonctionnalités de Hadoop

Consultez la liste suivante pour découvrir les nouvelles fonctionnalités Hadoop dans Amazon EMR 7.2.0.

- Support de S3 Glacier pour lire les objets restaurés : avec Amazon EMR version 7.2.0 et ultérieure, vous pouvez lire les objets Glacier restaurés depuis l'emplacement S3 de la table contenant le protocole. `S3A` Cette opération de lecture ignore les fichiers Glacier archivés s'ils sont toujours en cours de restauration.
- Support des écritures simultanées avec le commutateur magique S3A — Amazon EMR 7.2.0 introduit la prise en charge des écritures simultanées avec le validateur magique S3A. Cette fonctionnalité a été ajoutée à la communauté OSS Hadoop dans le cadre de HADOOP-18797.

Amazon EMR 6.6.0 – Notes de mise à jour de Hadoop

Amazon EMR 6.6.0 – Modifications apportées à Hadoop

Type	Description
Bogue	Correction d'enregistrements dupliqués lors de la lecture de fichiers BZip2 texte.
Rétroportage	HADOOP-18136 : vérifier la gestion par `FileUtils.unTar()` des fichiers .tar manquants
Rétroportage	HADOOP-17627 : rétroportage vers branch-3. 2 HADOOP-17371, HADOOP-17621,

Type	Description
	HADOOP-17625 pour mettre à jour Jetty vers la version 9.4.39
Rétroportage	HADOOP-17655 : mise à niveau de Jetty vers la version 9.4.40
Rétroportage	HADOOP-17796 : mise à niveau de jetty vers la version 9.4.43
Rétroportage	HADOOP-17661 : mvn versions:set ne parvient pas à analyser le fichier pom.xml
Rétroportage	HADOOP-17236 : augmenter snakeyaml à 1,26 pour atténuer la vulnérabilité CVE-2017-18640
Rétroportage	HADOOP-16717 : Supprimer GenericsUtil la dépendance de IsLog4JLogger sur Log4 jLoggerAdapter
Rétroportage	HADOOP-17633 : Passez json-smart à 2.4.2 et à 9.8 en raison de nimbus-jose-jwt CVEs
Rétroportage	HADOOP-17844 : mise à niveau de JSON smart vers la version 2.4.7
Rétroportage	HADOOP-17972 : rétroportage de HADOOP-17 683 (mise à jour de commons-io vers la version 2.8.0) pour la branche 3.2
Rétroportage	HADOOP-16555 : mise à jour de commons-c ompress vers la version 1.19
Rétroportage	HADOOP-17370 : mise à niveau de commons-compress vers la version 1.21
Rétroportage	HADOOP-17096 : correction ZStandard du décalage de la mémoire tampon d'entrée du compresseur

Type	Description
Rétroportage	HADOOP-17112 : les espaces blancs ne sont pas autorisés dans les chemins lors de l'enregis trement de fichiers sur s3a via un validateur
Rétroportage	HADOOP-13500 : Synchronisation de l'itération de l'objet Propriétés de configuration
Rétroportage	HDFS-14099 : Descripteur de trame inconnu lors de la décompression de plusieurs images dans le décompresseur ZStandard
Rétroportage	HDFS-16410 : Analyse XML non sécurisée OfflineEditsXmlLoader
Rétroportage	HDFS-14498 : LeaseManager peut tourner en boucle indéfiniment sur le fichier dont la création a échoué
Rétroportage	HDFS-15290 : entrée NPE au démarrage HttpServer NameNode
Rétroportage	HDFS-15293 : assouplir la condition d'accepta tion d'une image lors de la réception d'un point de contrôle
Rétroportage	HDFS-12979 : StandbyNode doit être téléchargé FsImage après le point de contrôle ObserverNode
Rétroportage	YARN-10538 : ajouter des nœuds de remise en service à la liste des nœuds mis à jour renvoyés à l'AM
Rétroportage	YARN-10472 : Rétroportage du YARN-10314 (YarnClientrenvoie uniquement les fichiers jar clients ombrés) vers NoClassDefFoundError la branche WebSocketException 3.2

Type	Description
Rétroportage	YARN-9968 : Public Localizer se ferme dans `NodeManager` en raison d'une exception `NullPointerException`
Rétroportage	YARN-10651 : `CapacityScheduler` a échoué avec NPE dans `AbstractYarnScheduler.updateNodeResource()`
Rétroportage	YARN-9339 : la métrique des applications en attente est incorrecte après le déplacement de l'application vers une nouvelle file d'attente
Rétroportage	YARN-10438 : gestion d'un `containerId` null dans `ClientRMService#getContainerReport()`
Rétroportage	YARN-7266 : ATS 1.5 ne démarre pas si les fichiers `RollingLevelDb` sont corrompus ou manquants
Rétroportage	YARN-9063 : ATS 1.5 ne démarre pas si les fichiers `RollingLevelDb` sont corrompus ou manquants
Rétroportage	YARN-9848 : revenir sur YARN-4946 (RM ne doit pas considérer une application comme `COMPLETED` lorsque l'agrégation des journaux n'est pas dans un état terminal).

Apache HBase

HBaseest une base de données distribuée open source, non relationnelle, développée dans le cadre du projet Hadoop de l'Apache Software Foundation. HBase s'exécute sur le système de fichiers distribué Hadoop (HDFS) afin de fournir des fonctionnalités de base de données non relationnelles à l'écosystème Hadoop. HBase est inclus dans les versions 4.6.0 et ultérieures d'Amazon EMR.

HBase fonctionne parfaitement avec Hadoop, partage son système de fichiers et sert d'entrée et de sortie directes au MapReduce framework et au moteur d'exécution. HBase s'intègre également à Apache Hive, permettant des requêtes de type SQL sur des tables, des jointures avec HBase des tables basées sur Hive et la prise en charge de la connectivité de base de données Java (JDBC). Pour plus d'informations HBase, consultez Apache HBase et HBase la documentation sur le site Web d'Apache. Pour un exemple d'utilisation HBase avec Hive, consultez le billet de blog sur le AWS Big Data Combine NoSQL et massively parallel analytics using HBase Apache and Apache Hive on Amazon EMR.

Avec Amazon EMR, vous pouvez également sauvegarder vos HBase données directement HBase sur Amazon Simple Storage Service (Amazon S3) et les restaurer à partir d'une sauvegarde créée précédemment lors du lancement d'un cluster. HBase Amazon EMR offre d'autres options d'intégration avec Amazon S3 pour la persistance des données et la reprise après sinistre.

- HBase sur Amazon S3 - Avec Amazon EMR version 5.2.0 et versions ultérieures, vous pouvez l'utiliser sur Amazon HBase S3 pour stocker le répertoire HBase racine et les métadonnées d'un cluster directement sur Amazon S3. Vous pouvez par la suite démarrer un nouveau cluster, en le désignant à l'emplacement du répertoire racine dans Amazon S3. Un seul cluster à la fois peut utiliser l' HBase emplacement dans Amazon S3, à l'exception d'un cluster en lecture et réplication. Pour de plus amples informations, veuillez consulter HBase sur Amazon S3 (mode de stockage Amazon S3).
- HBase read-replicas - Amazon EMR version 5.7.0 et versions ultérieures avec Amazon HBase S3 prend en charge les clusters de lecture et de réplication. Un cluster réplica en lecture offre un accès en lecture seule aux fichiers et métadonnées de stockage d'un cluster principal pour des opérations en lecture seule. Pour de plus amples informations, veuillez consulter Utilisation d'un cluster réplica en lecture.
- HBase Instantanés : comme alternative à Amazon S3, avec EMR version 4.0 et ultérieure, vous pouvez créer des instantanés de HBase vos données directement HBase sur Amazon S3, puis récupérer les données à l'aide des instantanés. Pour de plus amples informations, veuillez consulter Utilisation de HBase snapshots.

> **Important**
>
> Pour le dimensionnement des HBase clusters Amazon EMR, nous vous déconseillons d'utiliser le dimensionnement géré ou le dimensionnement avec des politiques personnalisées pour les clusters. HBase

Le tableau suivant répertorie la version HBase incluse dans la dernière version de la série Amazon EMR 7.x, ainsi que les composants avec lesquels Amazon EMR est installé. HBase

Pour connaître la version des composants installés HBase dans cette version, consultez la version 7.9.0 Versions des composants.

HBase informations de version pour emr-7.9.0

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-7.9.0	HBase 2.6.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Le tableau suivant répertorie la version HBase incluse dans la dernière version de la série Amazon EMR 6.x, ainsi que les composants avec lesquels Amazon EMR est installé. HBase

Pour connaître la version des composants installés HBase dans cette version, consultez la version 6.15.0 Versions des composants.

HBase informations de version pour emr-6.15.0

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.15.0	HBase 2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

> **Note**
>
> Apache HBase HBCK2 est un outil opérationnel distinct pour réparer HBase les régions et les tables système. Dans Amazon EMR version 6.1.0 et versions ultérieures, le fichier hbase-hbck2.jar est fourni `/usr/lib/hbase-operator-tools/` sur le nœud principal. Pour plus d'informations sur la création et l'utilisation de l'outil, consultez HBaseHBCK2.

Le tableau suivant répertorie la version HBase incluse dans la dernière version de la série Amazon EMR 5.x, ainsi que les composants avec lesquels Amazon EMR est installé. HBase

Pour la version des composants installés HBase dans cette version, voir la version 5.36.2 Versions des composants.

HBase informations de version pour emr-5.36.2

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.36.2	HBase 1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Rubriques

- Création d'un cluster avec HBase
- HBase sur Amazon S3 (mode de stockage Amazon S3)
- Journaux d'écriture anticipée (WAL) pour Amazon EMR
- Utilisation de la HBase coque
- Accédez aux HBase tables avec Hive
- Utilisation de HBase snapshots
- Configuration HBase
- Afficher l'interface HBase utilisateur
- Afficher les fichiers HBase journaux
- Moniteur HBase avec Ganglia

- Surveillance de l'EMR HBase avec Amazon CloudWatch
- Migration depuis les versions précédentes HBase
- HBase historique des versions

Création d'un cluster avec HBase

Les procédures décrites dans cette section décrivent les principes de base du lancement d'un cluster à l'aide du AWS Management Console et du AWS CLI. Pour des informations détaillées sur la planification, la configuration et le lancement de clusters Amazon EMR, consultez la section Planifier et configurer des clusters dans le guide de gestion Amazon EMR.

Création d'un cluster à HBase l'aide de la console

Pour connaître les étapes rapides permettant de lancer des clusters avec la console, consultez la section Premiers pas avec Amazon EMR dans le Guide de gestion Amazon EMR.

Pour lancer un cluster HBase installé à l'aide de la console

1. Ouvrez la console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com
2. Choisissez Create a cluster (Créer un cluster) et Go to advanced options (Accéder aux options avancées).
3. Pour Configuration des logiciels, choisissez une Version d'Amazon 4.6.0 ou ultérieure (nous recommandons la dernière version). Choisissez HBaseet d'autres applications comme vous le souhaitez.
4. Avec Amazon EMR version 5.2.0 et versions ultérieures, sous Paramètres de HBase stockage, sélectionnez HDFS ou S3. Pour de plus amples informations, veuillez consulter HBase sur Amazon S3 (mode de stockage Amazon S3).
5. Sélectionnez d'autres options si nécessaire, puis choisissez Create cluster (Créer le cluster).

Création d'un cluster à HBase l'aide du AWS CLI

Utilisez la commande suivante pour créer un cluster avec HBase installé :

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=HBase --use-default-roles --ec2-attributes KeyName=myKey \
```

```
--instance-type m5.xlarge --instance-count 3
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

Si vous l'utilisez HBase sur Amazon S3, spécifiez l'`--configurations`option en faisant référence à un objet de configuration JSON. L'objet de configuration doit contenir une `hbase-site` classification qui spécifie l'emplacement dans Amazon S3 où les HBase données sont stockées à l'aide de la `hbase.rootdir` propriété. Il doit contenir une classification `hbase`qui spécifie `s3` à l'aide de la propriété `hbase.emr.storageMode`. L'exemple suivant présente un extrait de JSON avec ces paramètres de configuration.

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3"
        }
    }
]
```

Pour plus d'informations HBase sur Amazon S3, consultezHBase sur Amazon S3 (mode de stockage Amazon S3). Pour plus d'informations sur les classifications, consultez Configuration des applications.

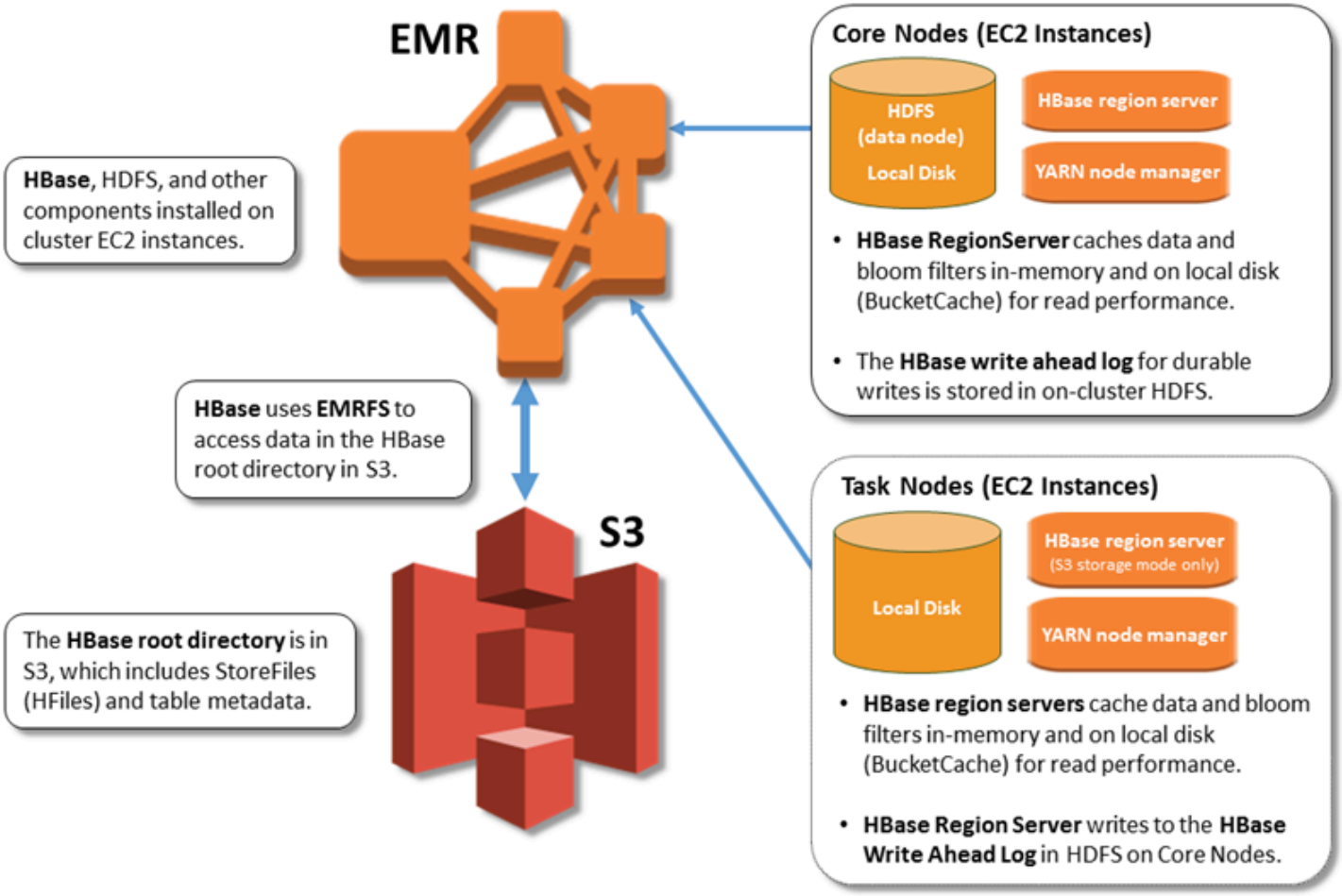
HBase sur Amazon S3 (mode de stockage Amazon S3)

Lorsque vous exécutez Amazon EMR version 5.2.0 ou ultérieure, vous pouvez l'activer HBase sur Amazon HBase S3, ce qui offre les avantages suivants :

- Le répertoire HBase racine est stocké dans Amazon S3, y compris les fichiers de HBase stockage et les métadonnées des tables. Ces données sont persistantes en dehors du cluster, disponibles dans toutes les zones de EC2 disponibilité Amazon, et vous n'avez pas besoin de les récupérer à l'aide de snapshots ou d'autres méthodes.
- Avec des fichiers de stockage dans Amazon S3, vous pouvez dimensionner votre cluster Amazon EMR plutôt en fonction de vos besoins de calcul qu'en fonction de vos besoins de données, avec une réplication x3 dans HDFS.
- En utilisant une version 5.7.0 ou ultérieure d'Amazon EMR, vous pouvez définir un cluster de réplica en lecture, ce qui vous permet de conserver des copies en lecture seule de données dans Amazon S3. Vous pouvez accéder aux données à partir du cluster de réplica en lecture pour réaliser des opérations de lecture simultanément et, dans le cas ou le cluster principal devient indisponible.
- Dans les versions 6.2.0 à 7.3.0 d'Amazon EMR, le HFile suivi persistant utilise une table HBase système appelée `hbase:storefile` pour suivre directement les HFile chemins utilisés pour les opérations de lecture. Cette fonctionnalité est activée par défaut et ne nécessite pas de migration manuelle. Dans les versions supérieures à 7.3.0, les HFile chemins sont suivis à l'aide d'un outil de suivi de fichiers, qui stocke HFile les chemins directement dans un méta-fichier, dans le répertoire du magasin.

> **Note**
>
> Pour les utilisateurs qui utilisent une version d'Amazon EMR antérieure à la version 7.4.0 et qui migrent vers EMR-7.4.0 ou version ultérieure, consultez la section Migration depuis les HBase versions précédentes et suivez la documentation de mise à niveau disponible pour garantir une transition fluide.

L'illustration suivante montre les HBase composants pertinents pour HBase Amazon S3.



Activation HBase sur Amazon S3

Vous pouvez l'activer HBase sur Amazon S3 à l'aide de la console Amazon EMR, de ou de l' AWS CLI API Amazon EMR. La configuration est une option lors de la création du cluster. Lorsque vous utilisez la console, vous choisissez le paramètre à l'aide des Advanced options (Options avancées). Lorsque vous utilisez l' AWS CLI, utilisez l'option `--configurations` pour fournir un objet de configuration JSON. Les propriétés de l'objet de configuration précisent le mode de stockage et l'emplacement du répertoire racine dans Amazon S3. L'emplacement Amazon S3 que vous spécifiez doit être dans la même région que votre cluster Amazon EMR. Un seul cluster actif à la fois peut utiliser le même répertoire HBase racine dans Amazon S3. Pour les étapes de console et un exemple détaillé de création de cluster à l'aide du AWS CLI, voir. Création d'un cluster avec HBase Un exemple d'objet de configuration apparaît dans l'extrait JSON suivant.

```
{
  "Classification": "hbase-site",
```

```
  "Properties": {
    "hbase.rootdir": "s3://amzn-s3-demo-bucket/my-hbase-rootdir"}
},
{
  "Classification": "hbase",
  "Properties": {
  "hbase.emr.storageMode":"s3"
  }
}
```

> **Note**
>
> Si vous utilisez un compartiment Amazon S3 comme formulaire `rootdir` HBase, vous devez ajouter une barre oblique à la fin de l'URI Amazon S3. Par exemple, vous devez utiliser `"hbase.rootdir: s3://amzn-s3-demo-bucket/"`, au lieu de `"hbase.rootdir: s3://amzn-s3-demo-bucket"`, pour éviter les problèmes.

Utilisation d'un cluster réplica en lecture

Après avoir configuré un cluster principal à l'aide d' HBase Amazon S3, vous pouvez créer et configurer un cluster en lecture seule qui fournit un accès en lecture seule aux mêmes données que le cluster principal. Cela est utile lorsque vous avez besoin d'un accès simultané à des données de requête ou d'un accès ininterrompu si le cluster principal devient indisponible. La fonctionnalité de réplica en lecture est disponible avec Amazon EMR version 5.7.0 et ultérieure.

Le cluster principal et le cluster de réplica en lecture sont définis de la même façon avec une différence significative. Les deux pointent le même emplacement `hbase.rootdir`. Cependant, la classification `hbase` pour le cluster de réplica en lecture comprend la propriété `"hbase.emr.readreplica.enabled":"true"`.

Le cluster de répliques en lecture seule est conçu pour les opérations en lecture seule, et aucune action manuelle de compactage ou d'écriture ne doit être effectuée sur celui-ci. Pour les versions d'Amazon EMR antérieures à 7.4.0, il est recommandé de désactiver le compactage sur le cluster de lecture-réplique lorsque vous activez la fonctionnalité de lecture-réplique. Cette précaution est nécessaire car, lorsque la fonctionnalité de HFile suivi permanent est activée sur le cluster principal, il est possible que le cluster en lecture-réplique compresse les tables système, ce qui peut provoquer un FileNotFoundException dysfonctionnement du cluster principal. La désactivation du compactage

sur le cluster en lecture permet d'éviter les incohérences de données entre le cluster principal et le cluster en lecture.

Par exemple, compte tenu de la classification JSON du cluster principal, comme indiqué précédemment dans cette rubrique, la configuration d'un cluster en lecture et en réplique pour les versions EMR antérieures à 7.4.0 est la suivante :

```
{
  "Classification": "hbase-site",
  "Properties": {
    "hbase.rootdir": "s3://amzn-s3-demo-bucket/my-hbase-rootdir",
    "hbase.regionserver.compaction.enabled": "false"
  }
},
{
  "Classification": "hbase",
  "Properties": {
  "hbase.emr.storageMode":"s3",
  "hbase.emr.readreplica.enabled":"true"
  }
}
```

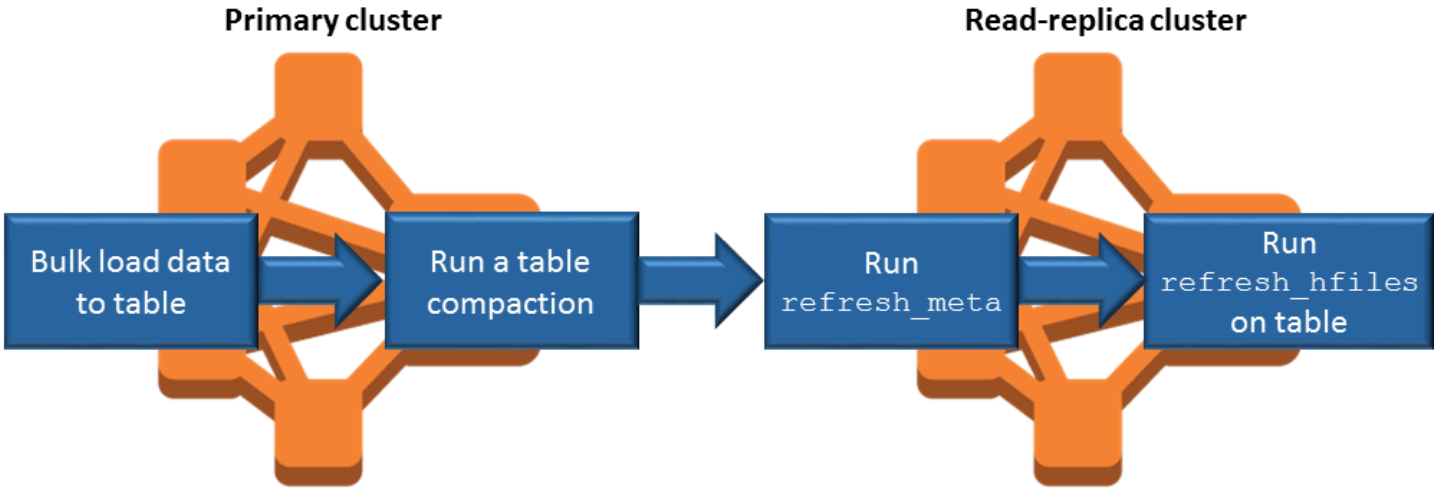
Pour les versions d'Amazon EMR ultérieures à 7.3.0, nous utilisons désormais Suivi des fichiers du magasin cette fonctionnalité. Il n'est donc pas nécessaire de désactiver les compactages.

Synchronisation de réplica en lecture lors d'ajout de données

Étant donné que la réplique en lecture utilise HBase StoreFiles et les métadonnées que le cluster principal écrit dans Amazon S3, la réplique en lecture est uniquement aussi actuelle que le magasin de données Amazon S3. Les conseils suivants vous aident à minimiser le retard entre le cluster principal et la réplica en lecture lorsque vous écrivez des données.

- Chargez des données en bloc sur le cluster principal chaque fois que possible. Pour plus d'informations, consultez la section Chargement groupé dans la HBase documentation d'Apache.
- Un vidage qui écrit des fichiers de stockage dans Amazon S3 devrait être réalisé dès que possible après l'ajout des données. Procédez au vidage manuel ou configurez les paramètres de vidage pour minimiser le temps de retard.

- Si les compactages doivent s'exécuter automatiquement, exécutez un compactage manuel afin d'éviter les incohérences lors du déclenchement des compactages.
- Sur le cluster de lecture et de réplication, lorsque des métadonnées ont changé, par exemple en cas de division de HBase région ou de compactage, ou lorsque des tables sont ajoutées ou supprimées, exécutez la commande. `refresh_meta`
- Sur le cluster de réplica en lecture, exécutez la commande `refresh_hfiles` lorsque des enregistrements sont ajoutés vers ou modifiés dans une table.



HFile Suivi permanent

HFile Le suivi permanent utilise une table HBase système appelée `hbase:storefile` pour suivre directement les HFile chemins utilisés pour les opérations de lecture. De nouveaux HFile chemins sont ajoutés à la table au fur et à mesure que des données supplémentaires sont ajoutées HBase. Cela supprime les opérations de renommage en tant que mécanisme de validation dans les HBase opérations de chemin d'écriture critiques et améliore le temps de restauration lors de l'ouverture d'une HBase région en lisant dans la table `hbase:storefile` système plutôt que dans la liste des répertoires du système de fichiers. Cette fonctionnalité est activée par défaut sur les versions 6.2.0 à 7.3.0 d'Amazon EMR et ne nécessite aucune étape de migration manuelle.

> **Note**
>
> Le HFile suivi permanent à l'aide de la table système HBase Storefile ne prend pas en charge la fonctionnalité de réplication des HBase régions. Pour plus d'informations sur HBase la

réplication des régions, consultez la section Lectures à haute disponibilité cohérentes avec la chronologie.

Désactiver le suivi permanent HFile

Le HFile suivi permanent est activé par défaut à partir de la version 6.2.0 d'Amazon EMR. Pour désactiver le HFile suivi permanent, spécifiez la modification de configuration suivante lors du lancement d'un cluster :

```
{
  "Classification": "hbase-site",
  "Properties": {
    "hbase.storefile.tracking.persist.enabled":"false",

 "hbase.hstore.engine.class":"org.apache.hadoop.hbase.regionserver.DefaultStoreEngine"
  }
}
```

> **Note**
>
> Lors de la reconfiguration du cluster Amazon EMR, tous les groupes d'instances doivent être mis à jour.

Synchronisation manuelle de la table Storefile

La table Storefile est mise à jour à mesure que de nouvelles HFile instances sont créées. Toutefois, si la table Storefile n'est plus synchronisée avec les fichiers de données pour une raison quelconque, les commandes suivantes peuvent être utilisées pour synchroniser manuellement les données :

Synchroniser la table Storefile dans une région en ligne :

```
hbase org.apache.hadoop.hbase.client.example.RefreshHFilesClient <table>
```

Synchroniser la table Storefile dans une région hors connexion :

- Supprimez la table Storefile znode.

```
echo "ls /hbase/storefile/loaded" | sudo -u hbase hbase zkcli
```

```
[<tableName>, hbase:namespace]
# The TableName exists in the list
echo "delete /hbase/storefile/loaded/<tableName>" | sudo -u hbase hbase zkcli
# Delete the Table ZNode
echo "ls /hbase/storefile/loaded" | sudo -u hbase hbase zkcli
[hbase:namespace]
```

- Attribuez la région (exécutez dans « hbase shell »).

```
hbase cli> assign '<region name>'
```

- Si l'assignation échoue.

```
hbase cli> disable '<table name>'
hbase cli> enable '<table name>'
```

Mise à l'échelle de la table Storefile

La table Storefile est divisée en quatre régions par défaut. Si la table Storefile est toujours soumise à une charge d'écriture importante, elle peut être divisée manuellement davantage.

Pour diviser une région chaude spécifique, utilisez la commande suivante (exécutée dans « hbase shell »).

```
hbase cli> split '<region name>'
```

Pour diviser la table, utilisez la commande suivante (exécutée dans « hbase shell »).

```
hbase cli> split 'hbase:storefile'
```

Suivi des fichiers du magasin

Par défaut, nous utilisons l'FileBasedStoreFileTrackerimplémentation. Cette implémentation crée de nouveaux fichiers directement dans le répertoire du magasin, évitant ainsi d'avoir à effectuer des opérations de renommage. Il conserve une liste des instances de hfile validées en mémoire, soutenue par des méta-fichiers dans chaque répertoire de magasin. Chaque fois qu'un nouveau hfile est validé, la liste des fichiers suivis dans le magasin concerné est mise à jour et un nouveau méta-fichier est écrit avec le contenu de la liste et en supprimant le méta-fichier précédent, qui contient une

liste obsolète. Pour plus d'informations sur le suivi des fichiers en magasin, consultez la section Suivi des fichiers en magasin dans le guide de HBase référence Apache.

L'implémentation du FileBasedStoreFile tracker est activée par défaut, à partir de la version 7.4.0 d'Amazon EMR :

```
{
  "Classification": "hbase-site",
  "Properties": {
    hbase.store.file-tracker.impl:
 "org.apache.hadoop.hbase.regionserver.storefiletracker.FileBasedStoreFileTracker"
}
```

Pour désactiver l' FileBasedStoreFileTracker implémentation, spécifiez la modification de configuration suivante lors du lancement d'un cluster :

```
{
  "Classification": "hbase-site",
  "Properties": {
    hbase.store.file-tracker.impl:
 "org.apache.hadoop.hbase.regionserver.storefiletracker.DefaultStoreFileTracker"
}
```

> **Note**
>
> Lors de la reconfiguration du cluster Amazon EMR, tous les groupes d'instances doivent être mis à jour.

Considérations opérationnelles

HBase les serveurs régionaux sont utilisés BlockCache pour stocker les lectures de données en mémoire et BucketCache pour stocker les lectures de données sur le disque local. En outre, les serveurs régionaux stockent les écritures de données en mémoire et utilisent les journaux d'écriture anticipée pour stocker les écritures de données dans HDFS avant que les données ne soient écrites dans Amazon HBase StoreFiles S3. MemStore Les performances de lecture de votre cluster dépendent de la fréquence à laquelle un enregistrement peut être récupéré à partir des caches en mémoire ou sur le disque. Une erreur de cache entraîne la lecture de l' StoreFile enregistrement depuis Amazon S3, ce qui présente une latence et un écart-type nettement supérieurs à ceux de la

lecture depuis HDFS. De plus, les taux de demandes maximal pour Amazon S3 sont inférieurs à ce qui peut être réalisé à partir du cache local, la mise en cache de données peut donc être importante pour les charges de travail avec une forte densité de lectures. Pour plus d'informations sur les performances d'Amazon S3, consultez la section Optimisation des performances dans le Guide de l'utilisateur Amazon Simple Storage Service.

Pour améliorer les performances, nous vous recommandons de mettre en cache la plus grande partie possible de votre ensemble de données dans le stockage d'EC2 instance. Comme il BucketCache utilise le stockage d' EC2 instance du serveur régional, vous pouvez choisir un type d' EC2 instance avec un stockage d'instance suffisant et ajouter du stockage Amazon EBS pour répondre à la taille de cache requise. Vous pouvez également augmenter la BucketCache taille des magasins d'instances attachés et des volumes EBS à l'aide de cette `hbase.bucketcache.size` propriété. Le paramètre par défaut est 8 192 MB.

En ce qui concerne les écritures, la fréquence des MemStore purges d'eau et le nombre de personnes StoreFiles présentes lors de compactages mineurs ou majeurs peuvent contribuer de manière significative à l'augmentation des temps de réponse des serveurs régionaux. Pour des performances optimales, pensez à augmenter la taille du MemStore purgeur et du multiplicateur de HRegion blocs, ce qui augmente le temps écoulé entre les compactages majeurs, mais augmente également le retard de cohérence si vous utilisez une réplique en lecture. Dans certains cas, vous pouvez obtenir de meilleures performances en utilisant des tailles de blocs de fichiers plus importantes (néanmoins inférieures à 5 Go) pour déclencher la fonction de chargement partitionné d'Amazon S3 dans EMRFS. La taille de bloc par défaut d'Amazon EMR est de 128 Mo. Pour de plus amples informations, veuillez consulter Configuration HDFS. Nous voyons rarement des clients dépasser une taille de bloc de 1 Go lorsque nous comparons les performances avec des vidages et des compactages. De plus, HBase les compactages et les serveurs régionaux fonctionnent de manière optimale lorsqu'il est StoreFiles nécessaire de compacter moins de serveurs.

Les tables peuvent mettre du temps pour arriver dans Amazon S3, car des répertoires volumineux doivent être renommés. Envisagez la désactivation des tables plutôt que leur suppression.

Il existe un HBase processus de nettoyage qui nettoie les anciens fichiers WAL et stocke les fichiers. Avec les version 5.17.0 et ultérieures d'Amazon EMR, l'appareil est activé dans le monde entier et les propriétés de configuration suivantes peuvent être utilisées pour contrôler un comportement plus propre.

Propriété de configuration	Valeur par défaut	Description

Propriété de configuration	Valeur par défaut	Description
`hbase.regionserver.hfilecleaner.large.thread.count`	1	Le nombre de threads alloués au nettoyage a expiré en grande quantité HFiles.
`hbase.regionserver.hfilecleaner.small.thread.count`	1	Le nombre de threads alloués au nettoyage a expiré petit HFiles.
`hbase.cleaner.scan.dir.concurrent.size`	Défini sur un quart de tous les noyaux disponibles.	Le nombre de threads à analyser dans les anciens WALs répertoires.
`hbase.oldwals.cleaner.thread.size`	2	Le nombre de threads à nettoyer WALs dans l'ancien WALs répertoire.

Avec les versions 5.17.0 et ultérieures d'Amazon EMR, l'opération de nettoyage peut voir un impact sur les performances de requêtes lors de l'exécution d'importantes charges de travail. Nous vous recommandons donc d'activer le nettoyeur uniquement pendant les heures creuses. Le nettoyeur possède les commandes HBase shell suivantes :

- `cleaner_chore_enabled` vérifie si l'appareil est activé.
- `cleaner_chore_run` exécute manuellement le nettoyeur pour supprimer les fichiers.
- `cleaner_chore_switch` active ou désactive le nettoyeur et le remet à son état précédent. Par exemple, `cleaner_chore_switch true` active le nettoyeur.

Propriétés pour HBase le réglage des performances sur Amazon S3

Les paramètres suivants peuvent être ajustés pour optimiser les performances de votre charge de travail lorsque vous l'utilisez HBase sur Amazon S3.

Propriété de configuration	Valeur par défaut	Description
`hbase.bucketcache.size`	8 192	La quantité d'espace disque, en Mo, réservée sur le serveur régional, sur les stockages d' EC2 instances Amazon et sur les volumes EBS pour le BucketCache stockage. Le paramètre s'applique à toutes les instances de serveur de la région. Des BucketCache tailles plus grandes correspondent généralement à de meilleures performances
`hbase.hregion.memstore.flush.size`	134217728	La limite des données, en octets, à laquelle le vidage d'un memstore vers Amazon S3 est déclenché.
`hbase.hregion.memstore.block.multiplier`	4	Un multiplicateur qui détermine la limite MemStore supérieure à laquelle les mises à jour sont bloquées. Si les MemStore dépassements `hbase.hregion.memstore.flush.size` sont multipliés par cette valeur, les mises à jour sont bloquées. MemStore des bouffées d'eau et du compactage peuvent survenir pour débloquer les mises à jour.
`hbase.hstore.blockingStoreFiles`	10	Le nombre maximum de ces informations peut exister dans un magasin avant StoreFiles

Propriété de configuration	Valeur par défaut	Description
		que les mises à jour ne soient bloquées.
`hbase.hregion.max.filesize`	10737418240	La taille maximale d'une région avant que la région ne soit fractionnée.

Arrêt et restauration d'un cluster sans perte de données

Pour arrêter un cluster Amazon EMR sans perdre de données qui n'ont pas été écrites sur Amazon S3, vous devez vider votre MemStore cache sur Amazon S3 pour y écrire de nouveaux fichiers de boutique. Tout d'abord, vous devez désactiver toutes les tables. La configuration de l'étape suivante peut être utilisée lorsque vous ajoutez une étape au cluster. Pour plus d'informations, consultez Utilisation des étapes à de l' AWS CLI et de la console dans le Guide de gestion d'Amazon EMR.

```
Name="Disable all tables",Jar="command-runner.jar",Args=["/bin/bash","/usr/lib/hbase/
bin/disable_all_tables.sh"]
```

Vous pouvez également exécuter la commande bash suivante directement.

```
bash /usr/lib/hbase/bin/disable_all_tables.sh
```

Après avoir désactivé toutes les tables, videz-les à l'`hbase:meta`aide du HBase shell et de la commande suivante.

```
flush 'hbase:meta'
```

Vous pouvez ensuite exécuter un script shell fourni sur le cluster Amazon EMR pour vider le cache. MemStore Vous pouvez l'ajouter soit comme une étape, soit en l'exécutant directement à l'aide d' AWS CLI sur le cluster. Le script désactive toutes les HBase tables, ce qui MemStore entraîne le transfert du serveur de chaque région vers Amazon S3. Si le script se termine avec succès, les données persistent dans Amazon S3 et le cluster peut être arrêté.

Pour redémarrer un cluster avec les mêmes HBase données, spécifiez le même emplacement Amazon S3 que le cluster précédent dans AWS Management Console ou à l'aide de la propriété `hbase.rootdir` de configuration.

Journaux d'écriture anticipée (WAL) pour Amazon EMR

Avec Amazon EMR 6.15 et versions ultérieures, vous pouvez écrire vos journaux d' HBase écriture anticipée (WAL) Apache sur le WAL Amazon EMR. Dans les versions inférieures d'Amazon EMR, lorsque vous créez un cluster avec l'option HBase sur Amazon S3, WAL est le seul HBase composant Apache stocké sur le disque local pour les clusters, et vous pouvez stocker d'autres composants tels que le répertoire racine, les fichiers de stockage (HFiles), les métadonnées des tables et les données sur Amazon S3.

Vous pouvez utiliser Amazon EMR WAL pour récupérer des données qui n'ont pas été transférées vers Amazon S3. Pour sauvegarder l'intégralité de vos HBase clusters, choisissez d'utiliser le service Amazon EMR WAL. En arrière-plan, `RegionServer` écrit vos journaux d' HBase écriture anticipée (WAL) dans le WAL pour Amazon EMR.

Si votre cluster ou l'AZ ne fonctionne pas correctement ou n'est plus disponible, vous pouvez créer un nouveau cluster, le pointer vers le même répertoire racine S3 et le même espace de travail Amazon EMR WAL, et récupérer automatiquement les données dans WAL en quelques minutes. Pour de plus amples informations, veuillez consulter Restauration depuis Amazon EMR WAL.

À partir des versions 7.3.0 et supérieures d'Amazon EMR, Amazon EMR crée plusieurs EMR pour chaque serveur et regroupe plusieurs régions HBase dans un seul Amazon WALs EMR WAL. Cela permet d'améliorer Apache HBase WAL afin d'améliorer l'utilisation des journaux et d'optimiser les coûts. Pour configurer le nombre d'instances Amazon EMR WAL par HBase `RegionServer`, utilisez le paramètre. `hbase.wal.regiongrouping.numgroups` Par défaut, ce paramètre est défini sur 2. Il existe deux tables système qui ne sont incluses dans aucun groupe WAL : meta et masterstore. Ces tables utilisent toujours leur propre individualité WALs.

Si vous utilisez une version inférieure à Amazon EMR 7.3.0, nous vous recommandons de désactiver manuellement les tables de l'ancien HBase cluster pour vous assurer que toutes les données du WAL Amazon EMR sont transférées vers Amazon S3. Supprimez ensuite l'ancien Amazon EMR WAL, mettez fin à l'ancien cluster et configurez un nouveau cluster qui exécute la dernière version. Si vous rencontrez des problèmes et que vous ne parvenez pas à désactiver les tables de l'ancien cluster, vous pouvez directement arrêter l'ancien cluster et `emr.wal.multiplex.migrate` définir sur`true`. sur le nouveau cluster. S'il est défini sur true, il HBase essaiera de rejouer les données des anciennes instances Amazon EMR WAL HBase lors de l'initialisation de la région et de supprimer les anciennes après la WALs rediffusion. Ce processus de rediffusion entraîne des coûts supplémentaires pour les lectures. Après la migration, nous vous recommandons de configurer

le cluster et de le `emr.wal.multiplex.migrate` définir sur`false`. Vous pouvez également supprimer le paramètre pour accélérer l'initialisation de HBase la région.

> **Note**
>
> Amazon EMR WAL supprime les données après HBase les avoir vidées. Si les données HBase ne sont pas supprimées, Amazon EMR WAL les conserve pendant une durée maximale de 30 jours. Après 30 jours, Amazon EMR WAL supprime automatiquement les données. Amazon EMR conserve les instances WAL jusqu'à 30 jours à compter de la résiliation d'un cluster EMR. Toutefois, si vous lancez un nouveau cluster compatible WAL depuis le même répertoire racine S3 au cours de ces 30 jours, Amazon EMR ne supprimera aucune des instances WAL de votre cluster précédent. Pour de plus amples informations, veuillez consulter Restauration depuis Amazon EMR WAL.

Les sections suivantes décrivent comment configurer et utiliser Amazon EMR WAL avec votre cluster EMR HBase activé.

Rubriques

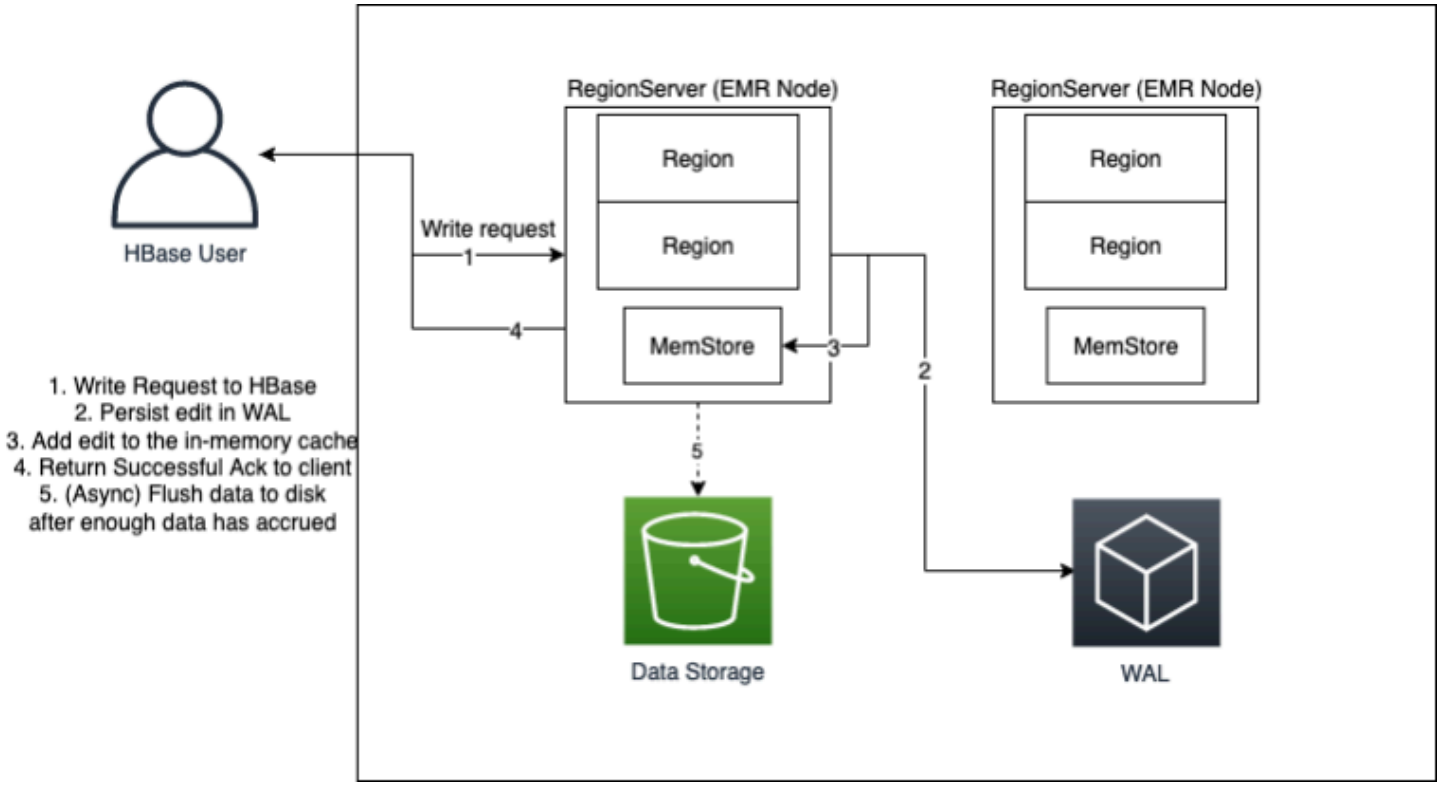
- Espaces de travail Amazon EMR WAL
- Autorisations requises pour Amazon EMR WAL
- Activation d'Amazon EMR WAL
- Restauration depuis Amazon EMR WAL
- Utilisation de configurations de sécurité avec Amazon EMR WAL
- Accédez à Amazon EMR WAL via AWS PrivateLink
- Comprendre la tarification et les statistiques d'Amazon EMR WAL
- Marquage des espaces de travail WAL
- Réplication entre clusters EMR WAL
- Considérations et régions relatives à Amazon EMR WAL
- Référence de la EMRWAL CLI Amazon EMR WAL ()

Espaces de travail Amazon EMR WAL

Amazon EMR WAL ajoute le concept d'espaces de travail WAL. Un espace de travail WAL est un conteneur logique de WALs. Chaque journal d'écriture anticipée dans Amazon EMR WAL est

encapsulé par un espace de travail WAL. Un cluster EMR écrit dans exactement un espace WALs de travail WAL que vous configurez au lancement du cluster, ou dans le `defaultWALworkspace` si vous ne spécifiez aucun espace de travail. Les espaces de travail WAL ne sont liés à aucune HBase terminologie existante telle que les espaces de noms.

Vous pouvez utiliser les espaces de travail WAL pour limiter les autorisations Amazon EMR WAL IAM afin d'inclure uniquement les espaces de travail auxquels le cluster doit accéder. Vous pouvez également étiqueter votre espace de travail WAL pour un contrôle d'accès basé sur des balises. Pour plus d'informations sur le balisage, consultezMarquage des espaces de travail WAL.



Autorisations requises pour Amazon EMR WAL

Pour que votre cluster puisse se connecter à Amazon EMR WAL, le profil d'instance du cluster nécessite certaines autorisations IAM :

- Amazon EMR WAL utilise le rôle AWSServiceRoleForEMRWALlié au service pour récupérer l'état d'un cluster. Amazon EMR crée automatiquement ce rôle lié au service lorsque vous créez un espace de travail WAL, ou HBase créera le rôle lié au service lorsque vous configurez un espace de travail pour Amazon EMR WAL et que le rôle lié au service n'existe pas encore.

Avant de pouvoir activer Amazon EMR WAL pour un cluster, vous devez configurer les autorisations afin de permettre la création automatique du rôle lié au AWSServiceRoleForEMRWAL service. Pour plus d'informations et un exemple d'instruction qui ajoute cette fonctionnalité, consultez la section Utilisation de rôles liés à un service pour la journalisation en écriture anticipée.

- Dans la mesure où Amazon EMR WAL utilise le HBase Write Ahead Log (WAL), vos clusters doivent utiliser HBase le WAL. Voici les autorisations IAM minimales dont vous avez besoin pour exécuter HBase. Ajoutez les éléments suivants à la politique d'autorisation de votre profil d'instance :

```
emrwal:DeleteWal
emrwal:CreateWal
emrwal:CreateWorkspace
emrwal:AppendEdit
emrwal:ReplayEdits
emrwal:GetCurrentWalTime
emrwal:CompleteWalFlush
emrwal:ListWALs
emrwal:DescribeWAL
emrwal:TrimWAL
emrwal:ArchiveWAL
emrwal:ArchiveWALCheckPoint
```

> **Note**
>
> Si vous limitez les autorisations pour Amazon EMR WAL à un ensemble minimal, certaines commandes EMRWALCLI ne disposeront pas des autorisations nécessaires pour s'exécuter.

Activation d'Amazon EMR WAL

Suivez les étapes suivantes pour activer l'écriture sur le WAL Amazon EMR lorsque vous créez un cluster avec le. AWS Command Line Interface

> **Note**
>
> Vous ne pouvez pas activer Amazon EMR WAL pour un cluster déjà en cours d'exécution, et vous ne pouvez pas lancer deux clusters avec le même répertoire racine S3. Pour de plus

amples informations, veuillez consulter Considérations et régions relatives à Amazon EMR WAL.

1. Avant de créer un cluster compatible Amazon EMR WAL, vous devez ajouter les autorisations requises au profil d'instance que vous prévoyez d'utiliser avec votre cluster. Pour de plus amples informations, veuillez consulter Autorisations requises pour Amazon EMR WAL.
2. Créez un cluster à partir du AWS CLI. Utilisez `--configurations` cette option pour fournir un objet de configuration JSON qui spécifie la `hbase.emr.wal.enabled` propriété, comme illustré dans l'exemple ci-dessous.

   - Spécifiez le mode de stockage et l'emplacement du répertoire racine dans Amazon S3. L'emplacement Amazon S3 que vous spécifiez doit se trouver dans la même région que votre cluster EMR, mais un seul cluster actif peut utiliser le même répertoire HBase racine dans S3 à la fois.
   - Créez votre cluster avec la configuration des groupes d'instances. Vous ne pouvez pas utiliser Amazon EMR WAL avec la configuration des flottes d'instances. Pour plus d'informations sur la création de clusters avec des groupes d'instances, consultez Configurer des groupes d'instances uniformes dans le guide de gestion Amazon EMR.
   - Pour connaître les étapes de création d'un cluster sur console, ainsi qu'un `create-cluster` exemple détaillé utilisant le AWS CLI, voir Création d'un cluster avec HBase.
3. Pour activer WAL pour le nouveau cluster, définissez la `hbase.emr.wal.enabled` propriété sur`true`. La commande suivante contient un extrait de code JSON avec un exemple d'objet de configuration.

```
aws emr create-cluster --name "hbasewal" --release-label emr-6.x.y \
--applications Name=HBase --use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m6i.xlarge --instance-count 1 --configurations hbase.json
$cat hbase.json
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
```

```
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

Lorsqu'il HBase est en ligne sur le cluster nouvellement créé, il écrit HBase automatiquement les données WAL sur le Amazon EMR WAL et utilise le Amazon EMR WAL à des fins de restauration.

Example 1 : Création d'un cluster EMR utilisant Amazon EMR WAL

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

Example 2 : Création d'un cluster EMR avec un espace de travail WAL personnalisé

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore",
            "emr.wal.workspace": "customWorkspaceName"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
```

```
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

Restauration depuis Amazon EMR WAL

Étant donné que le WAL Amazon EMR de votre cluster d'origine est conservé pendant 30 jours, vous pouvez restaurer et réutiliser le WAL pour un cluster nouvellement créé au cours de cette période de 30 jours. Lorsque vous lancez un nouveau cluster depuis le même répertoire racine S3, Amazon EMR conserve les instances WAL de votre ancien cluster. Si vous mettez fin à ce nouveau cluster, le délai de 30 jours redémarre à compter de la date de résiliation.

Utilisez la procédure suivante pour restaurer un WAL existant avec un nouveau cluster. Ce processus suppose que vous avez créé votre cluster d'origine avec Amazon EMR WAL activé.

1. Dans les 30 jours suivant la création d'un cluster compatible WAL, créez un nouveau cluster Région AWS identique au cluster d'origine. Le nouveau cluster peut se trouver dans la même AZ ou dans une autre AZ au sein de la même région que celle dans laquelle le cluster d'origine a été créé.

   Configurez les propriétés de l'objet pour spécifier le mode de stockage et l'emplacement du répertoire racine dans Amazon S3. L'emplacement Amazon S3 que vous spécifiez doit se trouver dans la même région que votre cluster EMR, mais un seul cluster actif peut utiliser le même répertoire HBase racine dans S3 à la fois.

   Pour connaître les étapes de création d'un cluster sur console, ainsi qu'un `create-cluster` exemple détaillé utilisant le AWS CLI, voir Création d'un cluster avec HBase.
2. Pour utiliser le WAL Amazon EMR existant pour le nouveau cluster, définissez la `hbase.emr.wal.enabled` propriété sur. `true` L'extrait de code JSON suivant montre un exemple d'objet de configuration.

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
```

```
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

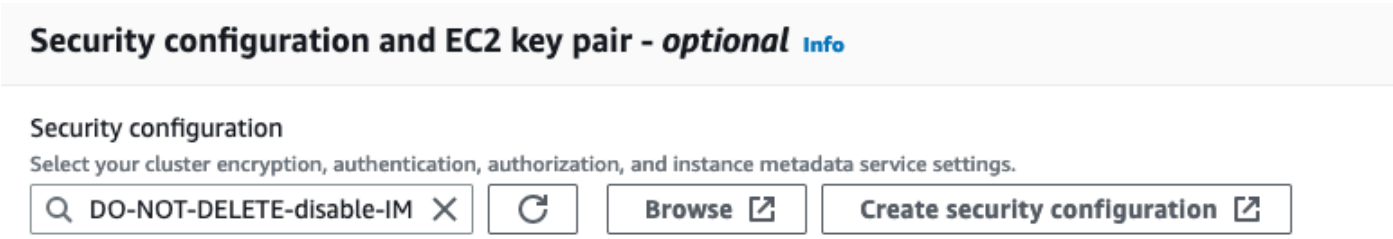
Utilisation de configurations de sécurité avec Amazon EMR WAL

Amazon EMR chiffre automatiquement les données en transit entre votre cluster et le service Amazon EMR WAL, ainsi que les données au repos dans Amazon EMR WAL. Pour plus d'informations, consultez Encryption at rest for Amazon EMR WAL. Vous pouvez également utiliser une configuration de sécurité pour importer vos propres clés depuis le service AWS Key Management Service (KMS) et chiffrer les données que vous stockez dans Amazon EMR WAL.

Utilisez l'une des méthodes suivantes pour sélectionner une configuration de sécurité lorsque vous créez un cluster :

Console

À partir de AWS Management Console, spécifiez la configuration sous Configuration de sécurité et paire de EC2 clés.



CLI

À partir de AWS CLI, définissez le `--security-configuration` paramètre lorsque vous utilisez la commande create-cluster.

Pour plus d'informations, consultez les sections Chiffrement au repos pour Amazon EMR WAL et Utiliser les configurations de sécurité pour configurer la sécurité du cluster dans le guide de gestion Amazon EMR.

Pour plus d'informations relatives à la sécurité sur WAL, consultez la section Utilisation de rôles liés à un service pour la journalisation en écriture anticipée.

Accédez à Amazon EMR WAL via AWS PrivateLink

Si vous souhaitez conserver votre connexion au sein du AWS réseau, Amazon EMR WAL propose une AWS PrivateLink assistance. Pour le configurer AWS PrivateLink, utilisez le AWS Management Console ou AWS Command Line Interface (AWS CLI) pour créer un point de terminaison VPC d'interface qui se connecte à Amazon EMR WAL. Pour plus d'informations, consultez la section Accès à un AWS service à l'aide d'un point de terminaison VPC d'interface dans le AWS PrivateLink Guide.

La procédure de base est la suivante :

1. Utilisez la console Amazon VPC pour créer un point de terminaison VPC. Sélectionnez Endpoints, puis Create endpoint.
2. Conservez la catégorie Service en tant que AWS services.
3. Dans la barre de recherche du panneau Services, tapez**`emrwal`**, puis sélectionnez le service intitulé`com.amazonaws.`*`region`*`.emrwal.prod`.
4. Sélectionnez votre VPC et enregistrez le point de terminaison. Assurez-vous d'associer les mêmes groupes de sécurité au point de terminaison VPC que vous attachez au cluster EMR.
5. Si vous le souhaitez, vous pouvez désormais activer les noms d'hôte DNS privés pour votre nouveau point de terminaison. Définissez Activer les noms d'hôte DNS et Activer le support DNS sur `true` pour votre VPC. Sélectionnez ensuite votre identifiant de point de terminaison, choisissez Modifier les paramètres VPC dans le menu Actions et activez les noms DNS privés.

   - Les noms d'hôte DNS privés du point de terminaison suivront le format`prod.emrwal.`*`region`*`.amazonaws.com`.
   - Si vous n'activez pas les noms d'hôte DNS privés, Amazon VPC vous fournit un nom de point de terminaison DNS au format suivant. *`endpointID`*`.prod.emrwal.`*`region`*`.vpce.amazonaws.com`
6. Pour utiliser votre AWS PrivateLink point de terminaison, modifiez la `emr.wal.client.endpoint` configuration lorsque vous créez votre cluster compatible Amazon EMR WAL, comme indiqué dans l'exemple suivant :

```
[
    {
```

```
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore",
            "emr.wal.workspace": "customWorkspaceName",
            "emr.wal.client.endpoint": "https://prod.emrwal.region.amazonaws.com"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

Vous pouvez également utiliser la politique VPCE pour autoriser ou restreindre l'accès au WAL Amazon APIs EMR. Pour plus d'informations, consultez la section Contrôler l'accès aux points de terminaison VPC à l'aide des politiques relatives aux points de terminaison dans le Guide.AWS PrivateLink

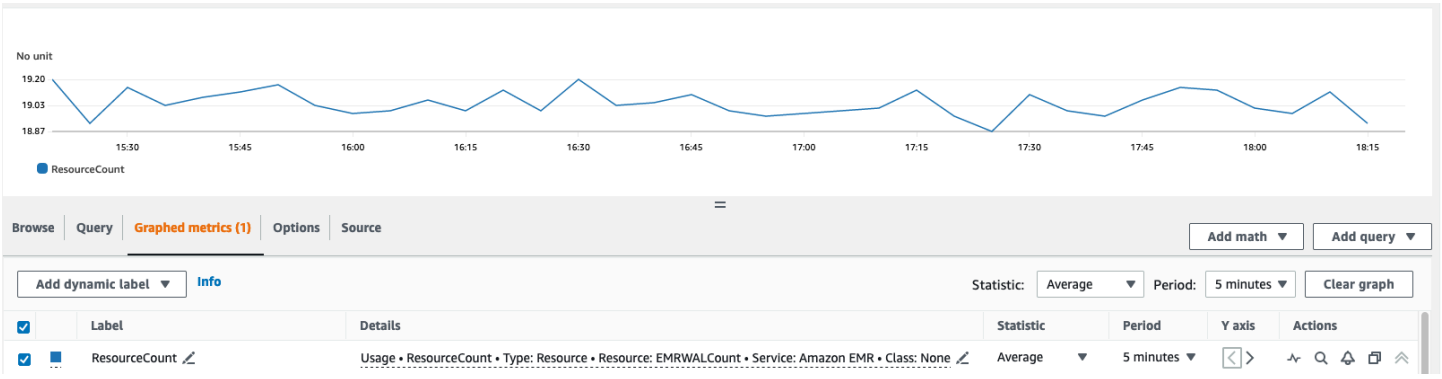
Comprendre la tarification et les statistiques d'Amazon EMR WAL

Unité de facturation des fonctionnalités de base	Détails
EMR-WAL-Read-GiB	API calls to read data from your table are billed as ReadRequestGiB. This includes `Get` and `Scan` operations. Reads are charged based on the sizes of the read items. Amazon EMR bills at a minimum of 1 byte. For example, if you read a 1234.12 bytes item, you're charged for 1235 bytes. Reads are aggregated every hour for billing and shown as GiBs.
EMR-WAL-Write-GiB	API calls to write data from your table are billed as Write-GiB. This includes `Put` operations. Writes are charged based on the sizes of the written items. Amazon EMR bills at a minimum of 1 byte. For example, if you write a 1234.12 bytes item, you're charged for 1235

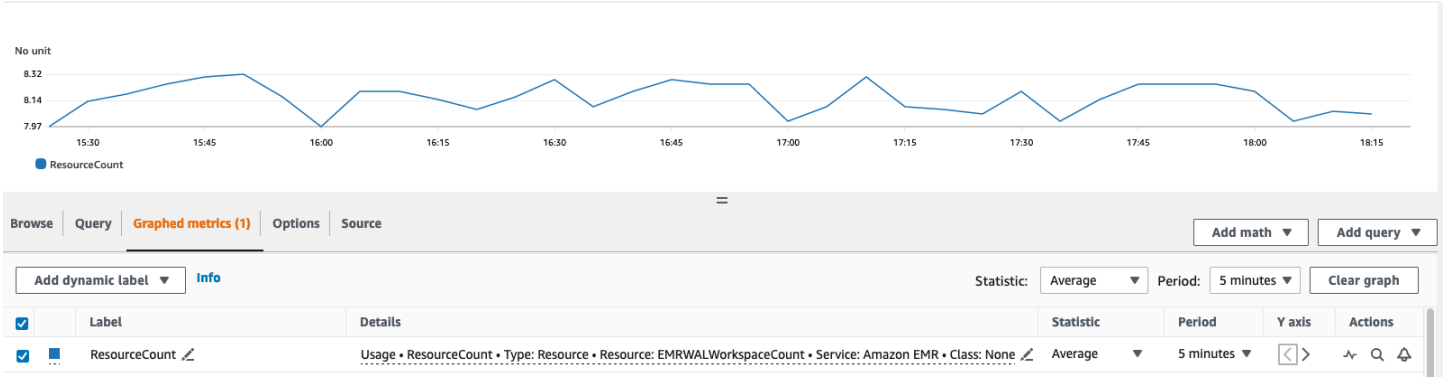
Unité de facturation des fonctionnalités de base	Détails
	bytes. Writes are aggregated every hour for billing and shown as GiBs.

EMR-WAL-WALHours	Le numéro WALs que vous enregistrez sur le service est facturé commeEMR-WAL-WALHours . Amazon EMR crée un WAL par région. HBase Par exemple, si vous créez 20 HBase tables, y compris des tables système, et que chaque table comporte deux HBase régions, vous utilisez 28 800 heures WAL, calculées comme suit :
	`20 tables`
	`x 2 Regions per table`
	`x 1 WAL per Region`
	`x 30 days`
	`x 24 hours`
	`----------`
	`28,800 EMR-WAL-WALHours`

Exemple **`EMRWALCount`** :



Exemple **`EMRWALWorkspaceCount`** :



Marquage des espaces de travail WAL

Vous pouvez ajouter des balises à un espace de travail lorsque vous créez un nouvel espace de travail et vous pouvez ajouter, supprimer ou répertorier des balises d'un espace de travail actif pour un cluster en cours d'exécution. Vous ne pouvez pas étiqueter les ressources individuelles de l'espace de travail, ni mettre à jour les balises existantes. Supprimez plutôt les balises indésirables de l'espace de travail et remplacez-les.

Vous pouvez baliser des espaces de travail à partir duEMRWAL CLI. Pour obtenir la liste des EMRWAL CLI commandes permettant de baliser les espaces de travail, consultez. Référence de la EMRWAL CLI Amazon EMR WAL ()

L'exemple de politique IAM suivant illustre un scénario qui autorise les opérations CRUDL de l'espace de travail uniquement avec la clé `resource_tag_allow_test_key` et la valeur de balisage appropriées : `resource_tag_allow_test_value`

```
{
  "Version": "2012-10-17",
  "Statement": [
      {
          "Action": [
              "EMRWAL:*"
          ],
          "Effect": "Allow",
          "Resource": [
              "*",
              "*"
          ],
          "Condition": {
              "StringEquals": {
                  "aws:ResourceTag/resource_tag_allow_test_key": [
```

```
                    "resource_tag_allow_test_value"
                ]
            }
        }
    }
  ]
}
```

Pour vérifier que la balise est désormais requise pour les opérations de l'espace de travail, utilisez la Référence de la EMRWAL CLI Amazon EMR WAL () `listTagsForResource`commande `tagAllowResourceTag` pour activer l'espace de travail avec la balise de ressource souhaitée. Si vous avez correctement configuré la condition, la commande aboutira.

```
emrwal listTagsForResource -r us-east-1 -arn arn:aws:emrwal:us-east-1:arn:workspace/
tagAllowResourceTag
Tag(Key=resource_tag_allow_test_key, Value=resource_tag_allow_test_value)
```

Réplication entre clusters EMR WAL

À partir d'EMR 7.5, EMR WAL prend en charge la HBase réplication entre clusters des journaux à écriture anticipée. Cette rubrique explique comment activer cette fonctionnalité et vérifier qu'elle fonctionne. Pour plus de détails sur la réplication en cluster, consultez la section Réplication en cluster dans la documentation Apache. HBase

> **Note**
>
> Des coûts de lecture supplémentaires sont associés aux journaux à écriture anticipée, car le processus de réplication lit les données depuis le WAL EMR local. Pour plus d'informations sur les coûts, consultez la section À propos des versions d'Amazon EMR.

Configuration de la réplication entre clusters

Afin d'offrir la même expérience utilisateur, l'activation de la fonctionnalité de réplication sur EMR WAL est la même que pour les journaux d' HBase écriture anticipée natifs. La procédure qui suit fournit un exemple simple. Pour plus d'informations, consultez la HBase documentation sur Cluster Replication.

1. Démarrez un cluster principal, qui est la source de réplication, avec EMR WAL activé. Pour activer les journaux à écriture anticipée, consultez. Activation d'Amazon EMR WAL Démarrez également un cluster homologue. Pour ce cluster de pairs, vous pouvez choisir d'activer ou non EMR WAL.
2. Sur les deux clusters, créez une table :

```
HBASE_CMD="sudo -u hbase hbase"
echo "create 'test_replication_table',{NAME => 'CF'}" | $HBASE_CMD shell
```

3. Ajoutez une configuration homologue sur le cluster principal et activez la réplication des tables. Lors de l'ajout du pair, il a besoin du nom d'hôte du nœud maître du cluster homologue, qui est PEER_DNS.

```
HBASE_CMD="sudo -u hbase hbase"
PEER_DNS="ip-10-1-1-0.ec2.com"
PEER_NAME="aws"
TABLE_NAME="test_replication_table"

## Create peering with the destination cluster
echo "add_peer '$PEER_NAME', CLUSTER_KEY => '$PEER_DNS:2181:/hbase'" | $HBASE_CMD
 shell

## List peers in the primary cluster to confirm peer setup
echo "list_peers" | $HBASE_CMD shell

## Enable table replication
echo "enable_table_replication '$TABLE_NAME'" | $HBASE_CMD shell
```

Confirmation de la réplication entre clusters

Après avoir effectué les étapes de configuration, la réplication est activée entre le cluster principal et le cluster homologue. Un test s'ensuit pour confirmer le bon fonctionnement de la réplication.

1. Ajoutez des données sur le cluster principal et vérifiez les données répliquées sur le cluster homologue.

```
## Write on primary cluster with HBase CLI

put 'test_replication_table', 'aaa', 'CF:a', 'aaa_a1'
put 'test_replication_table', 'bbb', 'CF:b', 'bbb_b1'
```

```
put 'test_replication_table', 'ccc', 'CF:c', 'ccc_c1'
```

2. Confirmez la réussite de la réplication sur le cluster homologue. Dans ce cas, vous devriez voir des données répliquées écrites du cluster principal vers le cluster homologue.

```
### Scan on peer cluster with HBase CLI

scan 'test_replication_table'
```

Considérations et régions relatives à Amazon EMR WAL

Considérations relatives à Amazon EMR WAL

La liste suivante décrit les principales considérations et limites d'Amazon EMR WAL :

- Amazon EMR WAL peut être utilisé avec les versions 6.15.0 et supérieures d'Amazon EMR.
- Amazon EMR WAL est un service payant optionnel. Vous payez pour ce que vous utilisez : lecture, écriture et stockage de données. Pour plus d'informations, consultez Comprendre la tarification et les statistiques d'Amazon EMR WAL la page de tarification d'Amazon EMR.
- Amazon EMR WAL utilise le HBase Write Ahead Log (WAL). Pour utiliser Amazon EMR WAL, vos clusters doivent utiliser HBase WAL.
- Pour activer Amazon EMR WAL lorsque vous créez un cluster, vous devez disposer des autorisations de rôle requises. Pour plus d'informations, voir Utilisation de rôles liés à un service pour la journalisation anticipée.
- Vous devez activer Amazon EMR WAL lorsque vous créez le cluster avec l'API AWS Management Console, ou AWS CLI, et vous devez utiliser la configuration des groupes d'instances. Vous ne pouvez pas activer Amazon EMR WAL dans un cluster en cours d'exécution si vous n'avez pas créé le cluster avec Amazon EMR WAL. Vous ne pouvez pas non plus modifier les `hbase-site` configurations pour activer Amazon EMR WAL dans un cluster en cours d'exécution.
- Vous ne pouvez activer Amazon EMR WAL que sur les clusters qui utilisent Amazon S3 comme répertoire racine.
- Avant la version 7.5.0 d'Amazon EMR, les enregistrements dans Amazon EMR WAL devaient être de 4 Mo ou moins. Mais avec Amazon EMR version 7.5.0 et versions ultérieures, la taille d'enregistrement maximale dans EMR WAL est configurable à l'aide de la propriété. `emr.wal.max.payload.size` La valeur par défaut est de 1 Go. L'exemple suivant définit la taille d'enregistrement maximale à 2 Go :

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "emr.wal.max.payload.size": "2147483648"
    }
  }
]
```

- Vous ne pouvez pas avoir plusieurs clusters actifs sur le même répertoire HBase racine dans Amazon S3.
- Vous ne pouvez pas activer Amazon EMR WAL sur des clusters de répliques en lecture.
- Le WAL est répliqué dans les zones de disponibilité du service géré.
- Le WAL survit au cluster et reste disponible pour le cluster suivant.
- Vous ne pouvez pas désactiver Amazon EMR WAL lors du lancement ou lorsque votre cluster est opérationnel (en cours d'exécution).
- Pour plus d'informations sur les limites du WAL et de l'espace de travail, consultez la section Points de terminaison et quotas Amazon EMR.

Disponibilité régionale pour Amazon EMR WAL

Le service Amazon EMR WAL est disponible dans les pays suivants : Régions AWS

- `ap-northeast-1` – Asie-Pacifique (Tokyo)
- `ap-northeast-2` – Asie-Pacifique (Séoul)
- `ap-southeast-1` – Asie-Pacifique (Singapour)
- `ap-south-1` – Asie-Pacifique (Mumbai)
- `ap-southeast-2` – Asie-Pacifique (Sydney)
- `eu-central-1` – Europe (Francfort)
- `eu-north-1` – Europe (Stockholm)
- `eu-west-1` – Europe (Irlande)
- `sa-east-1` – Amérique du Sud (São Paulo)
- `us-east-1` – USA Est (Virginie du Nord)

- `us-east-2` – USA Est (Ohio)
- `us-west-2` – USA Ouest (Oregon)

Les régions suivantes ne seront disponibles qu'avec Amazon EMR version 7.3.0 ou ultérieure :

- `ap-east-1` – Asie-Pacifique (Hong Kong)
- `af-south-1` – Afrique (Le Cap)
- `ca-central-1` – Canada (Centre)
- `eu-west-2` – Europe (Londres)

Référence de la EMRWAL CLI Amazon EMR WAL ()

L'interface de ligne de EMRWAL commande (EMRWAL CLI) est un outil unifié permettant de gérer votre journal d'écriture anticipée (WAL) pour Amazon EMR. Il est EMRWAL CLI livré avec des clusters EMR lorsque vous activez WAL au moment de créer un cluster. Pour plus d'informations sur l'activation de WAL, consultezJournaux d'écriture anticipée (WAL) pour Amazon EMR.

EMRWAL CLICela inclut les commandes suivantes :

Rubriques

- createWorkspace
- deleteWal
- deleteWorkspace
- listTagsForResource
- listWals
- listWorkspaces
- tagResource
- untagResource

`createWorkspace`

La `createWorkspace` commande crée un nouvel espace de travail Amazon EMR WAL.

Utilisation :

```
emrwal createWorkspace [-tags <tags>] [-e {endpoint}] [-r {Region}] -w {workspacename}
 [-h]
```

Exemple :

```
emrwal createWorkspace -w examplews
```

deleteWal

La `deleteWals` commande supprime le WAL Amazon EMR que vous spécifiez.

Utilisation :

```
emrwal deleteWal [-e {endpoint}] [-r {Region}] [-w {workspacename}] [-p <tablePrefix>]
 [-n <walName>] [-N <fullName>] [-R] [-m] [-h]
```

Exemple :

```
emrwal deleteWal -w examplews -p hbasetable -n examplewal
```

deleteWorkspace

La `deleteWorkspace` commande supprime l'espace de travail Amazon EMR WAL que vous spécifiez.

Utilisation :

```
emrwal deleteWorkspace [-e {endpoint}] [-r {Region}] -w {workspacename} [-h]
```

Exemple :

```
emrwal deleteWorkspace -w examplews
```

listTagsForResource

La `listTagsForResource` commande répertorie toutes les paires clé-valeur pour l'espace de travail Amazon EMR WAL que vous spécifiez.

Utilisation :

```
emrwal listTagsForResource -arn {resource-arn} [-e {endpoint}] [-r {Region}] [-h]
```

Exemple :

```
emrwal listTagsForResource -arn arn:aws:emrwal::1234567891234:workspace/examplews
```

`listWals`

La `listWals` commande répertorie tous les Amazon EMR présents WALs dans l'espace de travail que vous spécifiez.

Utilisation :

```
emrwal listWals [-nextToken {token-string}] [-pageSize {integer}] [-e {endpoint}] [-
r {Region}] [-w {workspacename}] [-p <tablePrefix>] [-M {integer}] [-h]
```

Exemple :

```
emrwal listWals -w examplews
```

`listWorkspaces`

La `listWorkspaces` commande répertorie tous les espaces de travail Amazon EMR WAL qui sont à votre disposition.

Utilisation :

```
emrwal listWorkspaces [-nextToken {token-string}] [-pageSize {integer}] [-e {endpoint}]
 [-r {Region}] [-M {integer}] [-h]
```

Exemple :

```
emrwal listWorkspaces
```

`tagResource`

La `tagResource` commande attribue une ou plusieurs paires clé-valeur à l'espace de travail Amazon EMR WAL que vous spécifiez.

Utilisation :

```
emrwal tagResource -arn {resource-arn} -tags <tags> [-e {endpoint}] [-r {Region}] [-h]
```

Exemple :

```
emrwal tagResource -arn arn:aws:emrwal::1234567891234:workspace/examplews -
tags tag_key=tag_value
```

`untagResource`

La `untagResource` commande annule l'attribution d'une ou de plusieurs paires clé-valeur à l'espace de travail Amazon EMR WAL que vous spécifiez.

Utilisation :

```
emrwal untagResource -arn {resource-arn} -tagKeys <tagKeys> [-e {endpoint}] [-
r {Region}] [-h]
```

Exemple :

```
emrwal untagResource -arn arn:aws:emrwal::1234567891234:workspace/examplews -
tagKeys tag_key
```

Utilisation de la HBase coque

Après avoir créé un HBase cluster, l'étape suivante consiste à vous HBase y connecter afin de pouvoir commencer à lire et à écrire des données (les écritures de données ne sont pas prises en charge sur un cluster en lecture-réplique). Vous pouvez utiliser le HBase shell pour tester les commandes.

Pour ouvrir la HBase coque

1. Utilisez SSH pour vous connecter au serveur principal du HBase cluster. Pour plus d'informations sur la façon de se connecter au nœud principal Amazon EMR via SSH, consultez la section Connexion au nœud principal via SSH dans le guide de gestion Amazon EMR.
2. Exécutez `hbase shell`. Le HBase shell s'ouvre avec un message similaire à ce qui suit.

   ```
   hbase(main):001:0>
   ```

Vous pouvez exécuter des commandes HBase shell à partir de l'invite. Pour plus d'informations sur les commandes shell et sur la façon de les appeler, tapez help à l' HBase invite et appuyez sur Entrée.

Créer une table

La commande suivante crée une table nommée « t1 », dotée d'une famille de colonnes unique nommée « f1 ».

```
hbase(main):001:0>create 't1', 'f1'
```

Placement d'une valeur

La commande suivante place la valeur « v1 » pour la ligne « r1 » dans la table « t1 » et la colonne « f1 ».

```
hbase(main):001:0>put 't1', 'r1', 'f1:col1', 'v1'
```

Obtention d'une valeur

La commande suivante obtient les valeurs pour la ligne « r1 » dans la table « t1 ».

```
hbase(main):001:0>get 't1', 'r1'
```

Supprimer une table

La commande suivante supprime et supprime le tableau « t1 ».

```
hbase(main):001:0>drop 'ns1:t1',false
```

La valeur booléenne indique si vous souhaitez archiver ou non votre table. Vous pouvez donc la définir sur `true` si vous souhaitez l'enregistrer. Vous pouvez également exécuter `drop 'ns1:t1'` sans booléen pour archiver la table.

Accédez aux HBase tables avec Hive

HBase et Apache Hive sont étroitement intégrés, ce qui vous permet d'exécuter des charges de travail de traitement massivement parallèles directement sur les données stockées dans. HBase Pour

utiliser Hive avec HBase, vous pouvez généralement les lancer sur le même cluster. Vous pouvez toutefois lancer Hive et HBase sur des clusters distincts. Exécuter HBase et Hive séparément sur différents clusters peut améliorer les performances, car cela permet à chaque application d'utiliser les ressources du cluster de manière plus efficace.

Les procédures suivantes montrent comment se connecter à un cluster HBase à l'aide de Hive.

> **Note**
>
> Vous ne pouvez connecter un cluster Hive qu'à un seul HBase cluster.

Pour connecter Hive à HBase

1. Créez des clusters séparés avec Hive et HBase installés ou créez un seul cluster avec les deux HBase et Hive installés.
2. Si vous utilisez des clusters distincts, modifiez vos groupes de sécurité afin que HBase les ports Hive soient ouverts entre ces deux nœuds principaux.
3. Utilisez SSH pour vous connecter au nœud principal du cluster sur lequel Hive est installé. Pour plus d'informations, consultez Connexion au nœud primaire à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
4. Lancez le shell Hive à l'aide de la commande suivante.

   ```
   hive
   ```

5. (Facultatif) Vous n'avez pas besoin de le faire si Hive HBase et Hive se trouvent sur le même cluster. Connectez le HBase client de votre cluster Hive au HBase cluster qui contient vos données. Dans l'exemple suivant, *public-DNS-name* est remplacé par le nom DNS public du nœud principal du HBase cluster, par exemple :`ec2-50-19-76-67.compute-1.amazonaws.com`.

   ```
   set hbase.zookeeper.quorum=public-DNS-name;
   ```

6. Exécutez les requêtes Hive sur vos HBase données comme vous le souhaitez ou consultez la procédure suivante.

Pour accéder aux HBase données depuis Hive

- Une fois la connexion établie entre le Hive et les HBase clusters (comme indiqué dans la procédure précédente), vous pouvez accéder aux données stockées sur le HBase cluster en créant une table externe dans Hive.

  L'exemple suivant, lorsqu'il est exécuté à partir de l'invite Hive sur le nœud principal, crée une table externe qui référence les données stockées dans une HBase table appeléeinputTable. Vous pouvez ensuite faire référence `inputTable` dans les instructions Hive pour interroger et modifier les données stockées dans le HBase cluster.

  ```
  set hbase.zookeeper.quorum=ec2-107-21-163-157.compute-1.amazonaws.com;

  create external table inputTable (key string, value string)
       stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
        with serdeproperties ("hbase.columns.mapping" = ":key,f1:col1")
        tblproperties ("hbase.table.name" = "t1");

  select count(key) from inputTable ;
  ```

Pour un cas d'utilisation plus avancé et un exemple de combinaison HBase avec Hive, consultez le billet de blog sur le AWS Big Data intitulé Combine NoSQL et massively parallel analytics using HBase Apache and Apache Hive on Amazon EMR.

Utilisation de HBase snapshots

HBase utilise une fonctionnalité de capture instantanée intégrée pour créer des sauvegardes légères de tables. Dans les clusters EMR, ces sauvegardes peuvent être exportées vers Amazon S3 à l'aide d'EMRFS. Vous pouvez créer un instantané sur le nœud principal à l'aide du HBase shell. Cette rubrique explique comment exécuter ces commandes de manière interactive avec le shell ou en effectuant une étape à l'`command-runner.jar`aide du AWS CLI ou AWS SDK pour Java. Pour plus d'informations sur les autres types de HBase sauvegardes, consultez la section HBase sauvegarde dans la HBase documentation.

Création d'un instantané à l'aide d'une table

```
hbase snapshot create -n snapshotName -t tableName
```

En utilisant le fichier command-runner.jar depuis AWS CLI :

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar",\
Args=[ "hbase", "snapshot", "create","-n","snapshotName","-t","tableName"]
```

AWS SDK pour Java

```
HadoopJarStepConfig hbaseSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("hbase","snapshot","create","-n","snapshotName","-t","tableName");
```

> **Note**
>
> Si votre nom d'instantané n'est pas unique, l'opération de création échoue avec un code de retour de `-1` ou `255`, mais vous ne verrez peut-être pas un message d'erreur indiquant ce qui s'est passé. Pour utiliser le même nom d'instantané, supprimez-le et recréez-le.

Suppression d’un instantané

```
hbase shell
>> delete_snapshot 'snapshotName'
```

Affichage des informations d'instantané

```
hbase snapshot info -snapshot snapshotName
```

Exportation d'un instantané vers Amazon S3

> **Important**
>
> Si vous ne spécifiez aucune `-mappers` valeur lors de l'exportation d'un instantané, HBase utilise un calcul arbitraire pour déterminer le nombre de mappeurs. Cette valeur peut être très grande, selon la taille de votre table, ce qui affecte de façon négative des tâches exécutées selon l'exportation. Pour cette raison, nous vous recommandons de spécifier le paramètre `-mappers`, le paramètre `-bandwidth` (qui spécifie la consommation de bande passante en mégaoctets par seconde), ou les deux, pour limiter les ressources de cluster utilisées par

l'opération d'exportation. Vous pouvez aussi exécuter l'opération d'instantané d'exportation durant une période de faible utilisation.

```
hbase snapshot export -snapshot snapshotName \
-copy-to s3://amzn-s3-demo-bucket/folder -mappers 2
```

En utilisant `command-runner.jar` à partir du AWS CLI :

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar",\
Args=[ "hbase", "snapshot", "export","-snapshot","snapshotName","-copy-to","s3://amzn-s3-demo-bucket/folder","-mappers","2","-bandwidth","50"]
```

AWS SDK pour Java:

```
HadoopJarStepConfig hbaseImportSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("hbase","snapshot","export",
      "-snapshot","snapshotName","-copy-to",
      "s3://bucketName/folder",
      "-mappers","2","-bandwidth","50");
```

Importer un instantané depuis Amazon S3

Bien qu'il s'agisse d'une importation, l' HBase option utilisée ici l'est toujours`export`.

```
sudo -u hbase hbase snapshot export \
-D hbase.rootdir=s3://amzn-s3-demo-bucket/folder \
-snapshot snapshotName \
-copy-to hdfs://masterPublicDNSName:8020/user/hbase \
-mappers 2
```

En utilisant `command-runner.jar` à partir du AWS CLI :

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar", \
Args=["sudo","-u","hbase","hbase snapshot export","-snapshot","snapshotName", \
"-D","hbase.rootdir=s3://amzn-s3-demo-bucket/folder", \
"-copy-to","hdfs://masterPublicDNSName:8020/user/hbase","-mappers","2","-chmod","700"]
```

AWS SDK pour Java:

```
HadoopJarStepConfig hbaseImportSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("sudo","-u","hbase","hbase","snapshot","export", "-D","hbase.rootdir=s3://
path/to/snapshot",
      "-snapshot","snapshotName","-copy-to",
      "hdfs://masterPublicDNSName:8020/user/hbase",
      "-mappers","2","-chuser","hbase");
```

Restaurer une table à partir d'instantanés dans le shell HBase

```
hbase shell
>> disable tableName
>> restore_snapshot snapshotName
>> enable tableName
```

HBase ne prend actuellement pas en charge toutes les commandes de capture d'écran présentes dans le HBase shell. Par exemple, aucune option de HBase ligne de commande ne permet de restaurer un instantané. Vous devez donc le restaurer dans un shell. Cela signifie que `command-runner.jar` doit exécuter une commande Bash.

> **Note**
>
> Comme la commande utilisée ici est `echo`, il est possible que votre commande shell échoue encore même si la commande exécutée par Amazon EMR retourne un code de sortie `0`. Vérifiez les journaux d'étape si vous choisissez d'exécuter une commande shell en tant qu'étape.

```
echo 'disable tableName; \
restore_snapshot snapshotName; \
enable tableName' | hbase shell
```

Voici l'étape utilisant l' AWS CLI. Commencez par créer le fichier `snapshot.json` suivant :

```
[
  {
    "Name": "restore",
```

```
    "Args": ["bash", "-c", "echo $'disable \"tableName\"; restore_snapshot
 \"snapshotName\"; enable \"tableName\"' | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file://./snapshot.json
```

AWS SDK pour Java:

```
HadoopJarStepConfig hbaseRestoreSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("bash","-c","echo $'disable \"tableName\"; restore_snapshot \"snapshotName
\"; enable \"snapshotName\"' | hbase shell");
```

Configuration HBase

Bien que les HBase paramètres par défaut devraient fonctionner pour la plupart des applications, vous pouvez modifier vos paramètres HBase de configuration. Pour ce faire, utilisez les propriétés des classifications de HBase configuration. Pour de plus amples informations, veuillez consulter Configuration des applications.

L'exemple suivant crée un cluster avec un répertoire HBase racine alternatif basé sur un fichier de configuration stocké dans Amazon S3. `myConfig.json`

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=HBase \
--instance-type m5.xlarge --instance-count 3 --configurations https://s3.amazonaws.com/
amzn-s3-demo-bucket/myfolder/myConfig.json
```

Le `myConfig.json` fichier spécifie la `hbase.rootdir` propriété de la classification de configuration `hbase-site`, comme illustré dans l'exemple suivant. *`ip-XXX-XX-XX-XXX.ec2.internal`*Remplacez-le par le nom d'hôte DNS interne du nœud principal du cluster.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.rootdir": "hdfs://ip-XXX-XX-XX-XXX.ec2.internal:8020/user/
myCustomHBaseDir"
    }
  }
]
```

> **Note**
>
> Avec la version 5.21.0 et ultérieures d'Amazon EMR, vous permet de remplacer les configurations de cluster et de spécifier des classifications de configuration supplémentaires pour chaque groupe d'instances dans un cluster en cours d'exécution. Pour ce faire, utilisez la console Amazon EMR, le AWS Command Line Interface (AWS CLI) ou le AWS SDK. Pour plus d'informations, consultez Fourniture d'une configuration pour un groupe d'instances dans un cluster en cours d'exécution.

Changements d'allocation de mémoire dans YARN

HBase ne s'exécute pas en tant qu'application YARN, il est donc nécessaire de recalculer la mémoire allouée à YARN et à ses applications, ce qui entraîne une réduction de la mémoire globale disponible pour YARN s'il HBase est installé. Vous devez en tenir compte lorsque vous planifiez de colocaliser des applications YARN HBase sur les mêmes clusters. Les types d'instance dotés de moins de 64 Go de mémoire disposent de la moitié de la mémoire disponibleNodeManager, qui est ensuite allouée au HBase RegionServer. Pour les types d'instance dotés d'une mémoire supérieure à 64 Go, HBase RegionServer la mémoire est limitée à 32 Go. En règle générale, la mémoire de réglage YARN est un multiple de la mémoire des tâches du MapReduce réducteur.

Les tableaux ci-dessous Valeurs par défaut des paramètres de configuration des tâches montrent les modifications apportées aux paramètres YARN en fonction de la mémoire requise pour HBase.

HBase numéros de port

Certains numéros de port choisis HBase sont différents de ceux par défaut. Les interfaces et les ports suivants sont destinés HBase à Amazon EMR.

HBase ports

utilisateur	Port	Protocole
HMaster	16000	TCP
HMaster UI	16010	HTTP
RegionServer	16020	TCP
RegionServer Infos	16030	HTTP
Serveur REST	8070	HTTP
Interface utilisateur de REST	8085	HTTP
Serveur Thrift	9090	TCP
Interface utilisateur du serveur Thrift	9095	HTTP

> **Important**
>
> Le `kms-http-port` est 9700 et `kms-admin-port` est 9701 dans Amazon EMR version 4.6.0 et ultérieure.

HBase paramètres du site pour optimiser

Vous pouvez définir tout ou partie des paramètres du HBase site afin d'optimiser le HBase cluster en fonction de la charge de travail de votre application. Nous recommandons les paramètres suivants comme point de départ dans votre enquête.

zookeeper.session.timeout

Le délai d'attente par défaut est de 40 secondes (40 000 ms). En cas de défaillance d'un serveur de région, il s'agit du temps nécessaire au serveur maître pour remarquer l'absence du serveur de région et commencer la récupération. Pour aider le serveur maître à récupérer plus rapidement, vous pouvez réduire cette valeur à une période de temps plus courte. L'exemple suivant utilise 30 secondes ou 30 000 ms :

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "zookeeper.session.timeout": "30000"
    }
  }
]
```

hbase.regionserver.handler.count

Cela définit le nombre de threads que le serveur de région laisse ouverts pour diffuser les demandes aux tables. La valeur par défaut de 10 est faible, afin d'empêcher les utilisateurs de supprimer leurs serveurs de région lorsqu'ils utilisent de grands tampons d'écriture avec un grand nombre de clients simultanés. La règle de base est de maintenir ce chiffre à un niveau bas lorsque la charge utile par demande se rapproche de l'ordre de Mo (gros achats, scans utilisant un cache volumineux) et élevé lorsque la charge utile est faible (get, small put ICVs, delete). L'exemple suivant augmente le nombre de threads ouverts à 30 :

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.regionserver.handler.count": "30"
    }
  }
]
```

hbase.hregion.max.filesize

Ce paramètre régit la taille, en octets, des régions individuelles. Par défaut, l’attribut est défini sur `1073741824`. Si vous écrivez beaucoup de données dans votre HBase cluster et que cela entraîne

des fractionnements fréquents, vous pouvez augmenter cette taille pour agrandir les différentes régions. Cela réduit les divisions mais augmente le temps nécessaire pour équilibrer la charge des régions d'un serveur à un autre.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.hregion.max.filesize": "1073741824"
    }
  }
]
```

hbase.hregion.memstore.flush.size

Ce paramètre régit la taille maximale de memstore, en octets, avant son vidage sur le disque. Par défaut, il s'agit de 134217728. Si votre charge de travail se compose de courtes périodes intensives d'opérations d'écriture, vous pouvez augmenter cette limite afin que toutes les écritures restent en mémoire au cours de ces périodes et soient vidées sur le disque plus tard. Cela peut améliorer les performances au cours de ces périodes intensives.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.hregion.memstore.flush.size": "134217728"
    }
  }
]
```

Considérations sur les performances

Activation de Z Garbage Collector (ZGC) pour HBase

Avec Amazon EMR version 7.10.0 et versions ultérieures, les utilisateurs peuvent facilement configurer les paramètres de collecte des déchets (GC) pour leur cluster. HBase Nous recommandons d'activer Z Garbage Collector (ZGC) pour obtenir une faible latence et des temps de pause GC inférieurs à la milliseconde.

Voici une configuration pour activer le ZGC pour HBase RegionServer (s) :

```
[
    {
        "Classification": "hbase-env",
        "Properties": {},
        "Configurations": [
            {
                "Classification": "export",
                "Properties": {
                    "JAVA_HOME": "/usr/lib/jvm/jre-21",
                    "HBASE_REGIONSERVER_GC_OPTS": "\"-XX:+UseZGC -XX:+ZGenerational\""
                }
            }
        ]
    }
]
```

> **Note**
>
> La variable d'environnement JAVA_HOME doit être définie lors de l'utilisation de Generational ZGC (recommandé) car elle a été introduite dans le JDK 21. Si vous souhaitez utiliser le mode non générationnel (sans **`-XX:+ZGenerational`**) de ZGC, il n'est pas nécessaire de définir le JAVA_HOME. Dans le JDK 24, le mode non générationnel de ZGC a été supprimé.

Réglage ZGC

1. Activer JVM Fixed Heap

   Le Z Garbage Collector (ZGC) fonctionne plus efficacement lorsqu'il est configuré avec de la mémoire fixe, éliminant ainsi le surcoût lié au retour de mémoire au système d'exploitation. Pour configurer la mémoire à segment fixe pour votre HBase cluster, utilisez la configuration suivante :

   ```
   [
       {
           "Classification": "hbase",
           "Properties": {
               "hbase.regionserver.fixed.heap.enabled": "true"
           }
       }
   ]
   ```

2. Activer le pré-toucher

   L'activation du précontact améliore les performances de collecte des déchets (GC) en réduisant la latence initiale et en fournissant des performances plus prévisibles. Pour activer le précontact pour votre HBase cluster, utilisez la configuration suivante :

```
[
    {
        "Classification": "hbase-env",
        "Properties": {},
        "Configurations": [
            {
                "Classification": "export",
                "Properties": {
                    "JAVA_HOME": "/usr/lib/jvm/jre-21",
                    "HBASE_REGIONSERVER_GC_OPTS": "\"-XX:+UseZGC -XX:+ZGenerational -
XX:+AlwaysPreTouch\""
                }
            }
        ]
    }
]
```

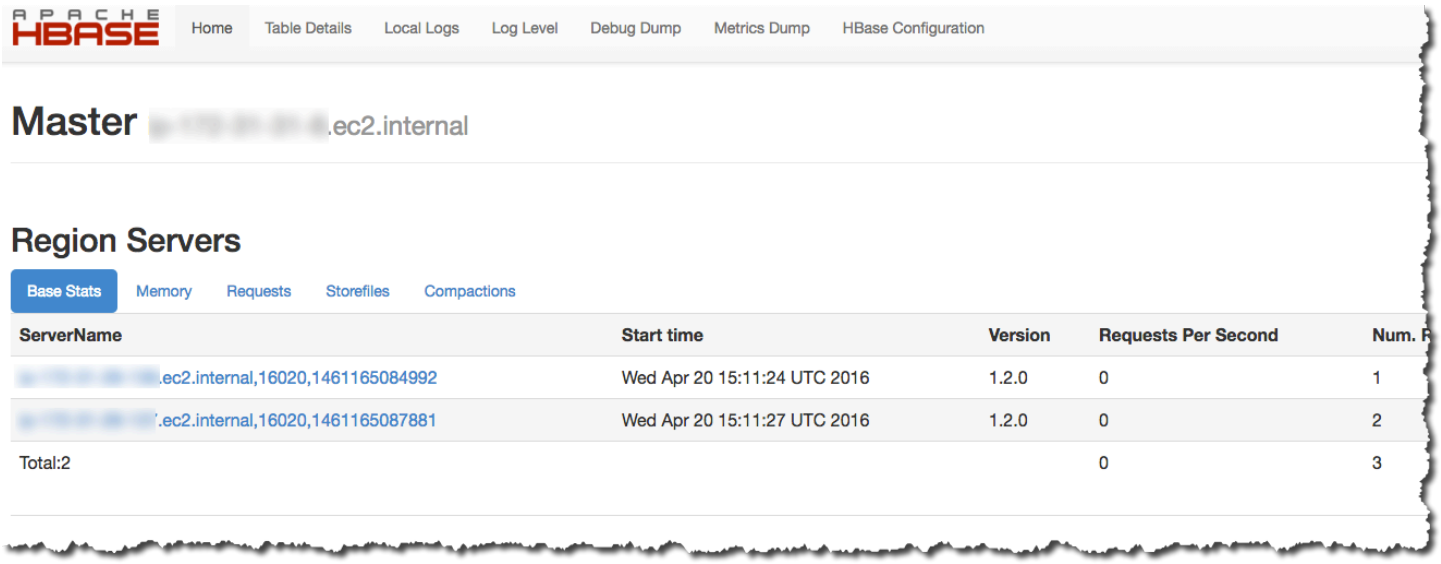
Afficher l'interface HBase utilisateur

> **Note**
>
> L'interface HBase utilisateur utilise des connexions HTTP non sécurisées par défaut. Pour activer le protocole HTTP sécurisé (HTTPS), définissez la `hbase.ssl.enabled` propriété de `hbase-site` classification sur `true` dans votre HBase configuration. Pour plus d'informations sur l'utilisation du protocole HTTP sécurisé (HTTPS) pour l'interface utilisateur HBase Web, consultez le guide de HBase référence Apache.

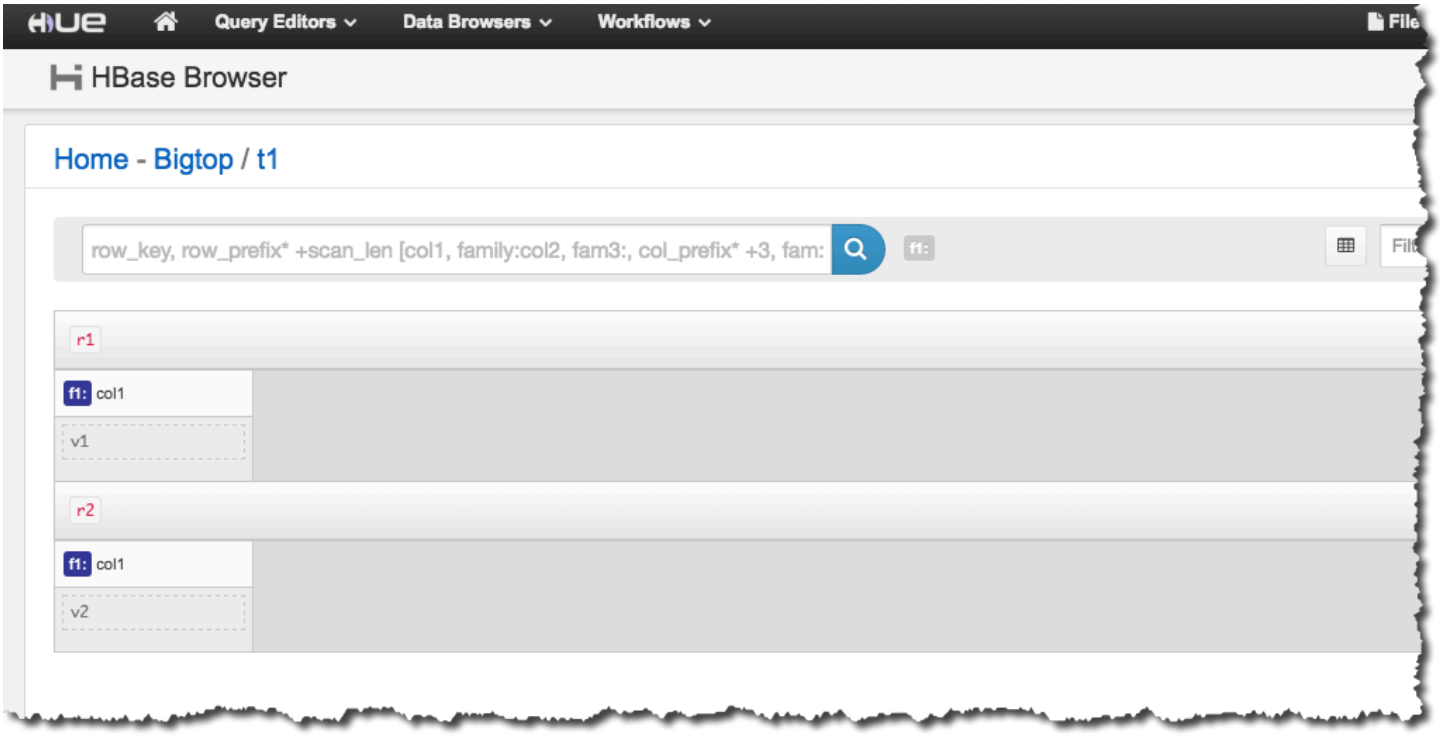
HBase fournit une interface utilisateur Web que vous pouvez utiliser pour surveiller votre HBase cluster. Lorsque vous exécutez HBase Amazon EMR, l'interface Web s'exécute sur le nœud principal et peut être consultée à l'aide de la redirection de port, également appelée création d'un tunnel SSH.

Pour afficher l'interface HBase utilisateur

1. Utilisez SSH pour créer un tunnel vers le nœud principal et créer une connexion sécurisée. Pour plus d'informations, consultez l'option 2, partie 1 : configurer un tunnel SSH vers le nœud principal à l'aide de la redirection de port dynamique dans le guide de gestion Amazon EMR.
2. Installez un navigateur Web avec un outil proxy, tel que le FoxyProxy plug-in pour Firefox, afin de créer un proxy SOCKS pour les AWS domaines. Pour plus d'informations, consultez l'option 2, partie 2 : configurer les paramètres du proxy pour afficher les sites Web hébergés sur le nœud principal dans le guide de gestion Amazon EMR.
3. Une fois le proxy défini et la connexion SSH ouverte, vous pouvez consulter l' HBase interface utilisateur en ouvrant une fenêtre de navigateur avec http : //:16010/master-status***master-public-dns-name***, où se trouve *master-public-dns-name* l'adresse DNS publique du nœud principal du cluster.



Vous pouvez également le voir HBase à Hue. Par exemple, l'illustration suivante montre la table, `t1`, que vous avez créée dans Utilisation de la HBase coque :



Pour plus d'informations sur Hue, consultez Hue.

Afficher les fichiers HBase journaux

Dans le cadre de son fonctionnement, HBase écrit des fichiers journaux contenant des détails sur les paramètres de configuration, les actions du démon et les exceptions. Ces fichiers journaux peuvent être utiles pour le débogage des problèmes HBase ainsi que pour le suivi des performances.

Si vous configurez votre cluster pour conserver les fichiers journaux dans Amazon S3, vous devez savoir que les journaux sont écrits dans Amazon S3 toutes les cinq minutes, si bien qu'il peut y avoir un léger retard avant que les derniers fichiers journaux soient disponibles.

Pour afficher HBase les journaux sur le nœud principal

- Vous pouvez consulter les HBase journaux actuels en utilisant SSH pour vous connecter au nœud principal et en accédant au `/var/log/hbase` répertoire. Ces journaux ne sont pas disponibles une fois que le cluster a été arrêté sauf si vous activez la journalisation dans Amazon S3 lorsque le cluster est lancé.

Pour consulter HBase les journaux sur Amazon S3

- Pour accéder aux HBase journaux et aux autres journaux de cluster sur Amazon S3, et pour qu'ils soient disponibles après la fin du cluster, spécifiez un compartiment Amazon S3 pour recevoir ces journaux lorsque vous créez le cluster. Pour cela, utilisez l'option `--log-uri`. Pour plus d'informations sur l'activation de la journalisation pour votre cluster, consultez la section Configuration de la journalisation et du débogage (facultatif) dans le Guide de gestion Amazon EMR.

Moniteur HBase avec Ganglia

> **Note**
>
> La dernière version d'Amazon EMR à inclure Ganglia était Amazon EMR 6.15.0. Pour surveiller votre cluster, les versions supérieures à 6.15.0 incluent l'agent Amazon CloudWatch .

Le projet open source Ganglia est un système distribué évolutif, conçu pour surveiller les clusters et les grilles tout en réduisant au maximum l'impact sur leurs performances. Lorsque vous activez Ganglia sur votre cluster, vous pouvez générer des rapports et afficher la performance du cluster dans son ensemble, ainsi qu'inspecter la performance des instances de chaque nœud. Pour plus d'informations sur le projet open source Ganglia, consultez http://ganglia.info/. Pour plus d'informations sur l'utilisation de Ganglia avec les clusters Amazon EMR, consultez Ganglia.

Une fois le cluster lancé avec Ganglia configuré, vous pouvez accéder aux graphes et aux rapports Ganglia à l'aide de l'interface graphique exécutée sur le nœud principal.

Ganglia stocke les fichiers journaux sur le nœud principal du `/mnt/var/lib/ganglia/rrds/` répertoire. Les versions antérieures d'Amazon EMR peuvent stocker des fichiers journaux dans le répertoire `/var/log/ganglia/rrds/`.

Pour configurer un cluster pour Ganglia et à HBase l'aide du AWS CLI

- Utilisez une commande `create-cluster` similaire à la commande suivante :

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=HBase Name=Ganglia --use-default-roles \
```

```
--ec2-attributes KeyName=myKey --instance-type m5.xlarge \
--instance-count 3
```

> **Note**
>
> Si le rôle de service Amazon EMR par défaut et le profil d' EC2 instance Amazon n'existent pas, une erreur se produit. Utilisez la commande `aws emr create-default-roles` pour les créer, puis réessayez.

Pour plus d'informations sur l'utilisation des commandes, consultez Commandes Amazon EMR dans l'interface AWS CLI.

Pour afficher les HBase métriques dans l'interface Web de Ganglia

1. Utilisez SSH pour créer un tunnel vers le nœud principal et créer une connexion sécurisée. Pour plus d'informations, consultez l'option 2, partie 1 : configurer un tunnel SSH vers le nœud principal à l'aide de la redirection de port dynamique dans le guide de gestion Amazon EMR.
2. Installez un navigateur Web avec un outil proxy, tel que le FoxyProxy plug-in pour Firefox, afin de créer un proxy SOCKS pour les AWS domaines. Pour plus d'informations, consultez l'option 2, partie 2 : configurer les paramètres du proxy pour afficher les sites Web hébergés sur le nœud principal dans le guide de gestion Amazon EMR.
3. Une fois le proxy défini et la connexion SSH ouverte, vous pouvez consulter les métriques Ganglia en ouvrant une fenêtre de navigateur avec http ://*master-public-dns-name*/ganglia/, où se *master-public-dns-name* trouve l'adresse DNS publique du serveur principal du cluster. HBase

Pour afficher les fichiers journaux Ganglia sur le nœud principal

- Si le cluster est toujours en cours d'exécution, vous pouvez accéder aux fichiers journaux en utilisant SSH pour vous connecter au nœud principal et en accédant au `/mnt/var/lib/ganglia/rrds/` répertoire. Pour EMR 3.x, naviguez jusqu'au répertoire `/var/log/ganglia/rrds`. Pour plus d'informations, consultez Connexion au nœud primaire à l'aide de SSH dans le Guide de gestion d'Amazon EMR.

Pour afficher les fichiers journaux de Ganglia sur Amazon S3

- Les fichiers journaux Ganglia ne sont pas automatiquement écrits dans Amazon S3 même si vous activez la journalisation pour votre cluster. Pour afficher les fichiers journaux Ganglia sur Amazon S3, vous devez les pousser manuellement de `/mnt/var/lib/ganglia/rrds/` vers le compartiment S3.

Surveillance de l'EMR HBase avec Amazon CloudWatch

À partir d'EMR 7.0, Amazon EMR fournit l' CloudWatch agent Amazon pour envoyer des métriques à Prometheus ou à CloudWatch Prometheus, en remplacement du système de surveillance Ganglia. Vous trouverez plus de détails dans le guide de configuration d'Amazon CloudWatch Agent.

L' CloudWatch agent Amazon EMR 7.0 a fourni une intégration de base et sa configuration a nécessité des efforts manuels, tels que le téléchargement de composants et l'utilisation d'actions de démarrage. Cependant, à partir de la version 7.1 de l'EMR, le processus est simplifié. L'utilisation de l' CloudWatch agent Amazon dans EMR 7.1 permet de surveiller facilement les métriques au niveau du système et spécifiques à l'application, y compris celles pour. HBase En utilisant l'API de configuration EMR, vous pouvez rapidement configurer et personnaliser le processus de collecte des métriques, et choisir où envoyer les données, que ce soit à Amazon CloudWatch ou à Prometheus. Cette flexibilité vous permet de surveiller de près vos HBase clusters, afin de garantir leur bon fonctionnement et leur efficacité.

Utiliser l'API de configuration EMR pour gérer l'ensemble de la configuration, ce qui rend le processus beaucoup plus fluide. Dans EMR 7.1, l' CloudWatch agent prend en charge trois principaux types de métriques :

- Mesures du système : elles incluent des indicateurs clés des performances du système, tels que : l'utilisation du processeur, l'utilisation du disque, l'utilisation de la mémoire, les E/S réseau, les processus et l'utilisation du swap.
- Métriques du démon Hadoop : ces métriques concernent les différents composants de Hadoop, notamment : les métriques, les DataNode métriques, les métriques YARN et NameNode les NodeManager métriques YARN. ResourceManager
- HBase Métriques — Ces mesures fournissent des informations sur HBase les performances : métriques HBase principales, métriques du serveur HBase régional, métriques du serveur HBase REST et métriques du serveur HBase Thrift.

Using the AWS CLI

Toutes les métriques pour Hadoop et Hadoop HBase sont basées sur JMX, ce qui signifie qu'ils utilisent des extensions de gestion Java pour fournir des informations détaillées. Voici comment configurer Amazon CloudWatch Agent pour surveiller HBase :

- Reportez-vous à la section Conditions préalables avant de créer un cluster avec Amazon CloudWatch Agent. Utilisez une `create-cluster` commande similaire à l'exemple qui apparaît après cette liste.
- Reportez-vous aux configurations prises en charge pour un HBase cluster.
- Reportez-vous à l'exemple ci-dessous pour des exemples de configurations pour configurer HBase la surveillance. Reportez-vous aux exemples de configurations pour l'`--configuration`entrée.

```
aws emr create-cluster --name "HBase cluster with CloudWatch agent" \
--release-label emr-7.1.0 \
--applications Name=HBase Name=AmazonCloudWatchAgent \
--ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \
--configurations file://./configurations.json \
--instance-count 3 --use-default-roles
```

Pour plus d'informations sur les exemples de configurations JSON pour les métriques, voir Configurer les métriques.

Using the console

Pour créer un cluster avec CloudWatch l'agent Amazon depuis la console, procédez comme suit :

Création d'un cluster avec l' CloudWatch agent depuis la console

1. Connectez-vous au et ouvrez la AWS Management Console console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com
2. Choisissez Créer un cluster.
3. Sous Nom et applications, choisissez Amazon EMR 7.0.0 ou version ultérieure.
4. Sous Ensemble d'applications, sélectionnez HBase les applications supplémentaires que vous souhaitez installer sur votre cluster, et incluez CloudWatch un agent dans vos sélections.

5. Développez les paramètres du logiciel. Vous pouvez ensuite saisir la configuration directement à l'aide de JSON ou d'une syntaxe raccourcie présentée dans le texte ombré dans la console. Sinon, vous pouvez fournir un URI Amazon S3 pour un fichier avec un objet JSON `Configurations`. Pour plus d'informations sur les exemples de configurations JSON pour les métriques, voir Configurer les métriques.
6. Procédez à la création du cluster adapté à votre cas d'utilisation.

Configurer des métriques

Pour surveiller le HBase Master, vous pouvez configurer Amazon CloudWatch Agent pour collecter des métriques spécifiques.

1. Configuration des métriques HBase principales : pour surveiller les métriques HBase principales, vous pouvez configurer Amazon CloudWatch Agent pour collecter des métriques spécifiques. Voici un exemple de configuration pour suivre les activités du responsable des missions principales :

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-master-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=Master,sub=AssignmentManager":
 "AssignFailedCount,AssignSubmittedCount",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

   Dans cette configuration :

- Nous spécifiez le MBean (`Hadoop:service=HBase,name=Master,sub=AssignmentManager`) pour collecter des métriques telles que `AssignFailedCount` et`AssignSubmittedCount`.
- Nous avons défini l'intervalle pour collecter ces métriques toutes les 30 secondes (30 000 millisecondes).

2. Configuration des métriques des serveurs HBase régionaux : pour surveiller les serveurs HBase régionaux, configurez l' CloudWatch agent comme suit :

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-region-server-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=RegionServer,sub=IPC":
 "numActiveHandler,numActivePriorityHandler",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

Cette configuration :

- Surveille les gestionnaires actifs sur le serveur de région (`numActiveHandler`,`numActivePriorityHandler`).
- Utilise un intervalle de 30 secondes pour la collecte des métriques.

3. Configuration des métriques du serveur HBase REST — Pour surveiller l'interface HBase REST, vous pouvez utiliser la configuration suivante :

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
```

```
      {
        "Classification": "emr-hbase-rest-server-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=REST": "successfulPut,successfulScanCount",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

Dans cet exemple, l' CloudWatch agent collecte des mesures relatives aux opérations PUT réussies et compte le nombre de scans toutes les 30 secondes.

4. Configuration des métriques du serveur HBase Thrift — Pour surveiller le HBase Thrift Server, vous pouvez configurer les métriques avec une configuration telle que la suivante :

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-thrift-server-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=Thrift,sub=ThriftOne":
 "BatchGet_max,BatchGet_mean",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

Cette configuration permet de suivre les temps maximum et moyen des opérations GET par lots sur le serveur Thrift.

Utilisation de la destination des métriques

Avec EMR 7.1, vous avez la possibilité d'envoyer vos données métriques à Amazon ou à CloudWatch Amazon Managed Service for Prometheus. Ce choix vous permet de vous intégrer facilement aux différents outils de surveillance, en fonction de vos besoins.

Envoi de métriques à Amazon CloudWatch

Pour envoyer des métriques à CloudWatch, utilisez cette configuration :

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "cloudwatch"
    },
    "Configurations": []
  }
]
```

Envoi de métriques à Amazon Managed Service pour Prometheus

Si vous préférez utiliser Prometheus, définissez la destination et fournissez l'URL du point de terminaison :

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "prometheus",
      "prometheus_endpoint": "https://aps-workspaces.region.amazonaws.com/
workspaces/workspace_id/api/v1/remote_write"
    },
    "Configurations": []
  }
]
```

`region`Remplacez-le par votre AWS région et par votre `workspace_id` identifiant d'espace de travail Prometheus. Cette configuration dirige vos HBase métriques vers Prometheus en utilisant le point de terminaison spécifié.

Avec la configuration ci-dessus, vous pouvez voir les métriques ci-dessous sous l'onglet Surveillance.

Migration depuis les versions précédentes HBase

Pour migrer les données d'une HBase version précédente, consultez la section Mise à niveau, numéro de HBase version et compatibilité dans le guide de HBase référence Apache. Vous devrez peut-être porter une attention particulière aux exigences relatives à la mise à niveau à partir des versions antérieures à 1.0 deHBase.

Migration vers Amazon EMR version 7.4.0 ou ultérieure

> **Note**
>
> Suivez ces instructions si vous migrez d'une version EMR antérieure à 7.4.0 vers une version supérieure à 7.3.0.

Si vous utilisez actuellement une version EMR avec la fonctionnalité Store File Tracking d'Amazon activée, incluse dans les versions 6.2.0 à 7.3.0, et que vous souhaitez passer à une version avec OSS Store File Tracking, disponible sur les versions EMR ultérieures à 7.3.0, procédez comme suit :

1. Dans le cluster existant :
    1. Désactivez le `hbase:storefile` tableau.
    2. Déposez la `hbase:storefile` table.
    3. chasse d'eau`hbase:meta`.
    4. Attendez que les métadonnées soient mises à jour.
2. Dans le nouveau cluster :
    1. Définissez le même répertoire Amazon S3 que le répertoire racine.
    2. Démarrez le cluster avec l'`DefaultStoreFileTracker`implémentation :

```
{
  "Classification": "hbase-site",
  "Properties": {
    hbase.store.file-tracker.impl:
 "org.apache.hadoop.hbase.regionserver.storefiletracker.DefaultStoreFileTracker"
}
```

3. Au niveau de la famille de tables ou de colonnes, utilisez les commandes suivantes pour modifier le suivi des fichiers de stockage :

   a. Modifiez le Store File Tracker de la table ou de la famille de colonnes de tables :

   ```
   hbase> change_sft 't1','FILE'
   hbase> change_sft 't2','cf1','FILE'
   ```

   b. Modifiez tout le Store File Tracker de la table correspondant à l'expression régulière (regex) donnée :

   ```
   hbase> change_sft_all 't.*','FILE'
   hbase> change_sft_all 'ns:.*','FILE'
   hbase> change_sft_all 'ns:t.*','FILE'
   ```

HBase historique des versions

Le tableau suivant répertorie la version HBase incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

HBase informations sur la version

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-7.9.0	2.6.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hma

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
		ster, hbase-client, hbase-reg ion-server, hbase-rest-server, hbase-thrift-server, hbase-ope rator-tools, zookeeper-client, zookeeper-server
emr-7,8.0	2.6.1	emrfs, emr-ddb, emr-goodi es, emr-kinesis, emr-s3-di st-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datano de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-ma pred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, hbase-hma ster, hbase-client, hbase-reg ion-server, hbase-rest-server, hbase-thrift-server, hbase-ope rator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-7,7.0	2.6.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-7.6.0	2.6.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-7.5.0	2.5.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-7.4.0	2.5.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-7.3.0	2.4,17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-7.2.0	2.4,17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-5.36.2	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-7.1.0	2.4,17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-7.0.0	2.4,17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.15.0	2.4,17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.14.0	2.4,17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.13.0	2.4,17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.12.0	2.4,17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.11.1	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.11.0	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.10.1	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.10.0	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.9.1	2.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-6.9.0	2.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.8.1	2.4.12	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-6.8.0	2.4.12	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.7.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-5.36.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.36.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.6.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.35.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.5.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.4.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.3.1	2.2.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.3.0	2.2.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.2.1	2.2.6-amzn-0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.2.0	2.2.6-amzn-0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.1.1	2.2.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.1.0	2.2.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.0.1	2.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-6.0.0	2.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.34.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.33.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.33.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.32.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.32.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.31.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.31.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.30.2	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.30.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.30.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
mr-5.29.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.28.1	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.28.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.27.1	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.27.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.26.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.25.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.24.1	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.24.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.23.1	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.23.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.22.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.21.2	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.21.1	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.21.0	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.20.1	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.20.0	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.19.1	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.19.0	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.18.1	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.18.0	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.17.2	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.17.1	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.17.0	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.16.1	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.16.0	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.15.1	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.15.0	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.14.2	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.14.1	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.14.0	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.13.1	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.13.0	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.12.3	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.12.2	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.12.1	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.12.0	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.11.4	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.11.3	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.11.2	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.11.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.11.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.10.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.10.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.9.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.9.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.8.3	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.8.2	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.8.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.8.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.7.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.7.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.6.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.6.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.4	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.5.3	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.2	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.5.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.4.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.4.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.3.2	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.3.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.3.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.2.3	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.2.2	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.2.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.2.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.1.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.1.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.0.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.0.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.0.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-5.0.0	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.6	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-4.9.5	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.4	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-4.9.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-4.9.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.5	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-4.8.4	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-4.8.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-4.8.0	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.4	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-4.7.3	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.2	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-4.7.1	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.0	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	HBase Version	Composants installés avec HBase
emr-4.6.1	1.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.6.0	1.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Apache HCatalog

HCatalog est un outil qui vous permet d'accéder aux tables de métastore Hive dans Pig, Spark SQL et/ou dans des applications personnalisées. MapReduce HCatalog possède une interface REST et un client de ligne de commande qui vous permettent de créer des tables ou d'effectuer d'autres opérations. Vous écrivez ensuite vos applications pour accéder aux tables à l'aide de HCatalog bibliothèques. Pour plus d'informations, consultez la section Utilisation HCatalog. HCatalog est inclus dans la version 4.4.0 d'Amazon EMR et les versions ultérieures.

HCatalog sur Amazon EMR version 5.8.0 et ultérieure prend en charge l'utilisation du catalogue de données AWS Glue comme métastore pour Hive. Pour plus d'informations, consultez Utilisation du catalogue de données AWS Glue en tant que metastore pour Hive.

Le tableau suivant répertorie la version HCatalog incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants avec lesquels Amazon EMR installe. HCatalog

Pour connaître la version des composants installés HCatalog dans cette version, consultez les versions des composants de la version 6.15.0.

HCatalog informations sur la version pour emr-6.15.0

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.15.0	HCatalog 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Le tableau suivant répertorie la version HCatalog incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants avec lesquels Amazon EMR installe. HCatalog

Pour connaître la version des composants installés HCatalog dans cette version, consultez Versions des composants de la version 5.36.2.

HCatalog informations de version pour emr-5.36.2

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.36.2	HCatalog 2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Rubriques

- Création d'un cluster avec HCatalog
- En utilisant HCatalog
- Exemple : création d'une HCatalog table et écriture dans celle-ci avec Pig
- HCatalog historique des versions

Création d'un cluster avec HCatalog

Bien qu' HCatalog il soit inclus dans le projet Hive, vous devez l'installer en tant qu'application autonome.

Pour lancer un cluster avec une HCatalog installation à l'aide de la console

La procédure suivante crée un cluster avec HCatalog installé. Pour plus d'informations sur la création de clusters à l'aide de la console, y compris les Options avancées, consultez Planifier et configurer des clusters dans le Guide de gestion Amazon EMR.

1. Ouvrez la console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com
2. Choisissez Create cluster (Créer un cluster) pour utiliser la Quick Create (Création rapide).
3. Pour le champ Software Configuration (Configuration des logiciels), choisissez la Version Amazon emr-4.4.0 ou une version ultérieure.
4. Dans le champ Sélectionner les applications, sélectionnez Toutes les applications ou HCatalog.
5. Sélectionnez d'autres options si nécessaire, puis choisissez Create cluster (Créer le cluster).

Pour lancer un cluster à HCatalog l'aide du AWS CLI

- Créez le cluster à l'aide de la commande suivante :

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Cluster with Hcat" --release-label emr-7.9.0 \
--applications Name=HCatalog --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

En utilisant HCatalog

Vous pouvez l'utiliser HCatalog dans différentes applications qui utilisent le métastore Hive. Les exemples de cette section montrent comment créer une table et l'utiliser dans le contexte de Pig et Spark SQL.

Désactivation de l'écriture directe lors de l'utilisation HCatalog HStorer

Chaque fois qu'une application utilise HCatStorer pour écrire dans une HCatalog table stockée dans Amazon S3, désactivez la fonction d'écriture directe d'Amazon EMR. Par exemple, désactivez l'écriture directe si vous utilisez la `STORE` commande Pig ou lors de l'exécution de tâches Sqoop qui écrivent HCatalog des tables dans Amazon S3. Vous pouvez désactiver la fonctionnalité d'écriture directe en définissant `mapred.output.direct.NativeS3FileSystem` et les configurations `mapred.output.direct.EmrFileSystem` sur `false`. L'exemple suivant montre comment définir ces configurations à l'aide de Java.

```
Configuration conf = new Configuration();
conf.set("mapred.output.direct.NativeS3FileSystem", "false");
conf.set("mapred.output.direct.EmrFileSystem", "false");
```

Création d'une table à l'aide de la HCat CLI et utilisation de ces données dans Pig

Créez le script suivant, `impressions.q`, sur votre cluster :

```
CREATE EXTERNAL TABLE impressions (
    requestBeginTime string, adId string, impressionId string, referrer string,
    userAgent string, userCookie string, ip string
  )
  PARTITIONED BY (dt string)
  ROW FORMAT
    serde 'org.apache.hive.hcatalog.data.JsonSerDe'
    with serdeproperties ( 'paths'='requestBeginTime, adId, impressionId, referrer,
 userAgent, userCookie, ip' )
  LOCATION 's3://[your region].elasticmapreduce/samples/hive-ads/tables/impressions/';
ALTER TABLE impressions ADD PARTITION (dt='2009-04-13-08-05');
```

Exécutez le script à l'aide de la HCat CLI :

```
% hcat -f impressions.q
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-
log4j.properties
OK
Time taken: 4.001 seconds
OK
Time taken: 0.519 seconds
```

Ouvrez le shell Grunt et accédez aux données dans `impressions` :

```
% pig -useHCatalog -e "A = LOAD 'impressions' USING
 org.apache.hive.hcatalog.pig.HCatLoader();
B = LIMIT A 5;
dump B;"
<snip>
(1239610346000,m9nwdo67Nx6q2kI25qt5On7peICfUM,omkxkaRpNhGPDucAiBErSh1cs0MThC,cartoonnetwork.com
 (compatible; MSIE 7.0; Windows NT 6.0; FunWebProducts; GTB6; SLCC1; .NET CLR
 2.0.50727; Media Center PC
 5.0; .NET,wcVWWTascoPbGt6bdqDbuWTPPHgOPs,69.191.224.234,2009-04-13-08-05)
(1239611000000,NjriQjdODgWBKnkGJUP6GNTbDeK4An,AWtXPkfaWGOaNeL9OOsFU8Hcj6eLHt,cartoonnetwork.com
 (compatible; MSIE 7.0; Windows NT 5.1; GTB6; .NET CLR
 1.1.4322),OaMU1F2gE4CtADVHAbKjjRRks5kIgg,57.34.133.110,2009-04-13-08-05)
(1239610462000,Irpv3oiu0I5QNQiwSSTIshrLdo9cM1,i1LDq44LRSJF0hbmhB8Gk7k9gMWtBq,cartoonnetwork.com
 (compatible; MSIE 6.0; Windows NT 5.2; SV1; .NET CLR 1.1.4322;
 InfoPath.1),QSb3wkLR4JAIut4Uq6FNFQIR1rCVwU,42.174.193.253,2009-04-13-08-05)
(1239611007000,q2Awfnpe0JAvhInaIp0VGx9KTs0oPO,s3HvTflPB8JIE0IuM6hOEebWWpOtJV,cartoonnetwork.com
 (compatible; MSIE 6.0; Windows NT 5.2; SV1; .NET CLR 1.1.4322;
 InfoPath.1),QSb3wkLR4JAIut4Uq6FNFQIR1rCVwU,42.174.193.253,2009-04-13-08-05)
(1239610398000,c362vpAB0soPKGHRS43cj6TRwNeOGn,jeas5nXbQInGAgFB8jlkhnprN6cMw7,cartoonnetwork.com
 (compatible; MSIE 8.0; Windows NT 5.1; Trident/4.0; GTB6; .NET CLR
 1.1.4322),k96n5PnUmwHKfiUI0TFP0TNMfADgh9,51.131.29.87,2009-04-13-08-05)
7120 [main] INFO  org.apache.pig.Main  - Pig script completed in 7 seconds and 199
 milliseconds (7199 ms)
16/03/08 23:17:10 INFO pig.Main: Pig script completed in 7 seconds and 199 milliseconds
 (7199 ms)
```

Accès à la table à l'aide de Spark SQL

Cet exemple crée un Spark DataFrame à partir de la table créée dans le premier exemple et montre les 20 premières lignes :

```
% spark-shell --jars /usr/lib/hive-hcatalog/share/hcatalog/hive-hcatalog-core-1.0.0-
amzn-3.jar
<snip>
scala> val hiveContext = new org.apache.spark.sql.hive.HiveContext(sc);
scala> val df = hiveContext.sql("SELECT * FROM impressions")
scala> df.show()
<snip>
16/03/09 17:18:46 INFO DAGScheduler: ResultStage 0 (show at <console>:32) finished in
 10.702 s
```

```
16/03/09 17:18:46 INFO DAGScheduler: Job 0 finished: show at <console>:32, took
 10.839905 s
+----------------+-------------------+-------------------+-----------------
+-------------------+-------------------+-------------+---------------+
|requestbegintime|               adid|       impressionid|         referrer|
   useragent|         usercookie|            ip|             dt|
+----------------+-------------------+-------------------+-----------------
+-------------------+-------------------+-------------+---------------+
|   1239610346000|m9nwdo67Nx6q2kI25...|omkxkaRpNhGPDucAi...|cartoonnetwork.com|
Mozilla/4.0 (comp...|wcVWWTascoPbGt6bd...|69.191.224.234|2009-04-13-08-05|
|   1239611000000|NjriQjdODgWBKnkGJ...|AWtXPkfaWGOaNeL9O...|cartoonnetwork.com|
Mozilla/4.0 (comp...|OaMU1F2gE4CtADVHA...| 57.34.133.110|2009-04-13-08-05|
|   1239610462000|Irpv3oiu0I5QNQiwS...|i1LDq44LRSJF0hbmh...|cartoonnetwork.com|
Mozilla/4.0 (comp...|QSb3wkLR4JAIut4Uq...|42.174.193.253|2009-04-13-08-05|
|   1239611007000|q2Awfnpe0JAvhInaI...|s3HvTflPB8JIE0IuM...|cartoonnetwork.com|
Mozilla/4.0 (comp...|QSb3wkLR4JAIut4Uq...|42.174.193.253|2009-04-13-08-05|
|   1239610398000|c362vpAB0soPKGHRS...|jeas5nXbQInGAgFB8...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610600000|cjBTpruoaiEtqLuMX...|XwlohBSs8Ipxs1bRa...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610804000|Ms3eJHNAEItpxvimd...|4SIj4pGmgVLl625BD...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610872000|h5bccHX6wJReDi1jL...|EFAWIiBdVfnxwAMWP...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610365000|874NBpGmxNFfxEPKM...|xSvE4XtGbdtXPF2Lb...|cartoonnetwork.com|
Mozilla/5.0 (Maci...|eWDEVVUphlnRa273j...| 22.91.173.232|2009-04-13-08-05|
|   1239610348000|X8gISpUTSqh1A5reS...|TrFblGT99AgE75vuj...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610743000|kbKreLWB6QVueFrDm...|kVnxx9Ie2i3OLTxFj...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610812000|9lxOSRpEi3bmEeTCu...|1B2sff99AEIwSuLVV...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610876000|lijjmCf2kuxfBTnjL...|AjvufgUtakUFcsIM9...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610941000|t8t8trgjNRPIlmxuD...|agu2u2TCdqWP08rAA...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610490000|OGRLPVNGxiGgrCmWL...|mJg2raBUpPrC8OlUm...|       corriere.it|
Mozilla/4.0 (comp...|r2k96t1CNjSU9fJKN...|   71.124.66.3|2009-04-13-08-05|
|   1239610556000|OnJID12x0RXKPUgrD...|P7Pm2mPdW6wO8KA3R...|       corriere.it|
Mozilla/4.0 (comp...|r2k96t1CNjSU9fJKN...|   71.124.66.3|2009-04-13-08-05|
|   1239610373000|WflsvKIgOqfIE5KwR...|TJHd1VBspNcua0XPn...|       corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
|   1239610768000|4MJR0XxiVCU1ueXKV...|1OhGWmbvKf8ajoU8a...|       corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
```

```
|   1239610832000|gWIrpDiN57i3sHatv...|RNL4C7xPi3tdar2Uc...|       corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
|   1239610789000|pTne9k62kJ14QViXI...|RVxJVIQousjxUVI3r...|        pixnet.net|
Mozilla/5.0 (Maci...|1bGOKiBD2xmui9OkF...| 33.176.101.80|2009-04-13-08-05|
+---------------+--------------------+--------------------+-----------------
+--------------------+--------------------+-------------+---------------+
only showing top 20 rows


scala>
```

Exemple : création d'une HCatalog table et écriture dans celle-ci avec Pig

Vous pouvez créer une HCatalog table et utiliser Apache Pig pour y écrire par le biais de HCat Storer à l'aide d'une source de données dans Amazon S3. HCatalog nécessite que vous désactiviez l'écriture directe, sinon l'opération échoue silencieusement. Définissez les configurations `mapred.output.direct.NativeS3FileSystem` et `mapred.output.direct.EmrFileSystem` sur `false` à l'aide de la classification `mapred-site`, ou manuellement depuis le shell Grunt. L'exemple suivant illustre une table créée à l'aide de l' HCat interface de ligne de commande, suivie des commandes exécutées dans le shell Grunt pour remplir la table à partir d'un fichier d'exemple de données dans Amazon S3.

Pour exécuter cet exemple, il est nécessaire de se connecter au nœud principal à l'aide de SSH.

Créez un fichier de HCatalog script`wikicount.q`, avec le contenu suivant, qui crée une HCatalog table nommée`wikicount`.

```
CREATE EXTERNAL TABLE IF NOT EXISTS wikicount(
col1 string,
col2 bigint
)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\001'
STORED AS ORC
LOCATION 's3://amzn-s3-demo-bucket/hcat/wikicount';
```

Utilisez une commande HCat CLI pour exécuter le script à partir du fichier.

```
hcat -f wikicount.q
```

Ensuite, démarrez le shell Grunt avec l'option `-useHCatalog`, définissez des configurations pour désactiver l'écriture directe, chargez les données depuis un emplacement S3, puis écrivez les résultats dans la table wikicount.

```
pig -useHCatalog
SET mapred.output.direct.NativeS3FileSystem false;
SET mapred.output.direct.EmrFileSystem false;
A = LOAD 's3://support.elasticmapreduce/training/datasets/wikistats_tiny/' USING
 PigStorage(' ') AS (Site:chararray, page:chararray, views:int, total_bytes:long);
B = GROUP A BY Site;
C = FOREACH B GENERATE group as col1, COUNT(A) as col2;
STORE C INTO 'wikicount' USING org.apache.hive.hcatalog.pig.HCatStorer();
```

HCatalog historique des versions

Le tableau suivant répertorie la version HCatalog incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

HCatalog informations sur la version

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-7.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-s

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
		erver, hive-client, mariadb-s erver
emr-7.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver
emr-7,7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-7.6.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.5.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-7.4.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.3.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-7.2.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.36.2	2.3.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-7.1.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.0.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.15.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.14.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.13.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.12.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.11.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.11.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.10.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.10.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.9.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.8.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.36.1	2.3.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.36.0	2.3.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.6.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.35.0	2.3.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.5.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.4.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.3.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.3.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.2.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.2.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.1.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.1.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.0.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-6.0.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.34.0	2.3.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.33.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.33.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.32.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.32.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.31.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.31.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.30.2	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.30.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.30.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
mr-5.29.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.28.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.28.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.27.1	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.27.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.26.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.25.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.24.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.24.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.23.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.23.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.22.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.21.2	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.21.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.21.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.20.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.20.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.19.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.19.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.18.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.18.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.17.2	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.17.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.17.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.16.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.16.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.15.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.15.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.14.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.14.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.14.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.13.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.13.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.12.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.12.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.12.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.12.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.11.4	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.11.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.11.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.11.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.11.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.10.1	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.10.0	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.9.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.9.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.8.3	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.8.2	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.8.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.8.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.7.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.7.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.6.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.6.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.4	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.5.3	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.5.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.4.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.4.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.3.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.3.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.3.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.2.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.2.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.2.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.2.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.1.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.1.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.0.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.0.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.0.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-5.0.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.6	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-4.9.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-4.9.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-4.9.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-4.8.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-4.8.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-4.8.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-4.7.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-4.7.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-4.6.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server
emr-4.6.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server

Étiquette de version Amazon EMR	HCatalog Version	Composants installés avec HCatalog
emr-4.5.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server
emr-4.4.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server

Apache Hive

Hive est un entrepôt de données open source et un package analytique qui s'exécute au-dessus d'un cluster Hadoop. Les scripts Hive utilisent un langage de type SQL appelé Hive QL (langage de requête), qui extrait des modèles de programmation et prend en charge les interactions classiques d'entrepôt de données. Hive vous permet d'éviter les difficultés liées à l'écriture de jobs Tez basés sur des graphes acycliques dirigés (DAGs) ou des MapReduce programmes dans un langage informatique de niveau inférieur, tel que Java.

Hive étend le modèle SQL en incluant des formats de sérialisation. Vous pouvez également personnaliser le traitement d'une requête en créant un schéma de table correspondant à vos données, sans toucher au données elles-même. Alors que SQL prend uniquement en charge les types de valeur primitifs tels que les dates, les chiffres et les chaînes, les valeurs figurant dans les tableaux Hive sont des éléments structurés, tels que des objets JSON, tous types de données définis par l'utilisateur ou toutes fonctions écrites en Java.

Pour plus d'informations sur Hive, consultez le site http://hive.apache.org/.

Le tableau suivant répertorie la version de Hive incluse dans la dernière version d’Amazon EMR série 7.x, ainsi que les composants qu’Amazon EMR installe avec Hive.

Pour la version des composants installés avec Hive dans cette version, consultez la version 7.9.0 Versions des composants.

Informations sur la version de Hive pour emr-7.9.0

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-7.9.0	Hive 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-ya

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
		rn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Le tableau suivant répertorie la version de Hive incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Hive.

Pour connaître la version des composants installés avec Hive dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version de Hive pour emr-6.15.0

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.15.0	Hive 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Le tableau suivant répertorie la version de Hive incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Hive.

Pour la version des composants installés avec Hive dans cette version, voir la version 5.36.2 Versions des composants.

Informations sur la version Hive pour emr-5.36.2

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.36.2	Hive 2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

À partir d'Amazon EMR 5.18.0, vous pouvez utiliser le référentiel d'artefacts d'Amazon EMR pour générer le code de votre tâche en fonction des versions exactes des bibliothèques et des dépendances qui sont disponibles avec des versions spécifiques d'Amazon EMR. Pour de plus amples informations, veuillez consulter Vérification des dépendances à l'aide du référentiel d'artefacts d'Amazon EMR.

Rubriques

- Différences et considérations pour Hive sur Amazon EMR
- Configuration d'un metastore externe pour Hive
- Utilisation du pilote Hive JDBC
- Amélioration des performances de Hive

- Utilisation de Live Long and Process (LLAP) de Hive
- Chiffrement dans Hive
- Historique des versions de Hive

Différences et considérations pour Hive sur Amazon EMR

Différences entre Apache Hive sur Amazon EMR et Apache Hive

Cette section décrit les différences entre Hive sur Amazon EMR et les versions par défaut de Hive disponibles à http://svn.apache. org/viewvc/hive/branches/.

Autorisation Hive

Amazon EMR prend en charge l'autorisation Hive pour HDFS, mais pas pour EMRFS et Amazon S3. Les clusters Amazon EMR s'exécutent avec l'autorisation désactivée par défaut.

Comportement de fusion de fichiers Hive avec Amazon S3

Apache Hive fusionne les petits fichiers à la fin d'une tâche de type map-only si `hive.merge.mapfiles` est défini sur true et la fusion est déclenchée uniquement si la taille de sortie moyenne de la tâche est inférieure au paramètre `hive.merge.smallfiles.avgsize`. Le comportement d'Amazon EMR Hive est identique si le chemin de la sortie finale est sur HDFS. Si le chemin de sortie se trouve dans Amazon S3, le paramètre `hive.merge.smallfiles.avgsize` est ignoré. Dans ce cas, la tâche de fusion est toujours déclenchée si `hive.merge.mapfiles` est défini sur `true`.

Transactions ACID et Amazon S3

Amazon EMR 6.1.0 et versions ultérieures prennent en charge les transactions Hive ACID (Atomicity, Consistency, Isolation, Durability) afin de respecter les propriétés ACID d'une base de données. Grâce à cette fonctionnalité, vous pouvez exécuter des opérations INSERT, UPDATE, DELETE et MERGE dans des tables gérées par Hive, avec des données dans Amazon Simple Storage Service (Amazon S3).

Fonctionnalités LLAP Hive

La fonctionnalité LLAP ajoutée dans la version 2.0 d'Apache Hive par défaut n'est pas prise en charge dans la version Hive 2.1.0 sur Amazon EMR version 5.0.

Amazon EMR version 6.0.0 et ultérieure prend en charge la fonctionnalité Live Long and Process (LLAP) pour Hive. Pour de plus amples informations, veuillez consulter Utilisation de LLAP pour Hive.

Différences Hive entre les versions Amazon EMR 4.x et 5.x

Cette section indique les différences à prendre en compte avant de migrer une implémentation de Hive de la version Hive 1.0.0 sur la version 4.x d'Amazon EMR à la version Hive 2.x sur la version 5.x d'Amazon EMR.

Différences et considérations opérationnelles

- Ajout de la prise en charge des transactions ACID (atomicité, cohérence, isolation et durabilité) : cette différence entre Hive 1.0.0 sur Amazon EMR 4.x et Apache Hive par défaut a été supprimée.
- Suppression des écritures directes vers Amazon S3 : cette différence entre Hive 1.0.0 sur Amazon EMR et Apache Hive par défaut a été supprimée. Désormais, Hive 2.1.0 sur la version 5.x d'Amazon EMR crée, lit et écrit des fichiers temporaires stockés dans Amazon S3. Par conséquent, vous n'avez plus besoin de créer une table temporaire dans le système de fichiers HDFS local du cluster pour lire la même table et y écrire. Si vous utilisez des compartiments activés pour la gestion des versions, veillez à gérer ces fichiers temporaires comme indiqué ci-dessous.
- Gestion des fichiers temporaires lors de l'utilisation de compartiments Amazon S3 versionnés : lorsque vous exécutez des requêtes Hive pour lesquelles la destination des données générées est Amazon S3, de nombreux fichiers temporaires et répertoires sont créés. Il s'agit du nouveau comportement, comme indiqué précédemment. Si vous utilisez des compartiments S3 activés pour la gestion des versions, ces fichiers temporaires encombrement Amazon S3 et entraînent des coûts s'ils ne sont pas supprimés. Ajustez vos règles de cycle de vie afin que les données ayant un préfixe `/_tmp` soient supprimées après un délai court, par exemple cinq jours. Pour plus d'informations, consultez Configuration de cycle de vie.
- Mise à jour de Log4j par log4j 2 : si vous utilisez log4j, vous devrez peut-être modifier votre configuration de journalisation en raison de cette mise à niveau. Pour plus d'informations, consultez Apache log4j 2.

Différences et considérations en termes de performances

- Différences de performances avec Tez : dans la version 5.x d'Amazon EMR, Tez devient le moteur d'exécution par défaut pour Hive, au lieu de. MapReduce Tez offre des performances améliorées pour la plupart des flux de travail.

- Tables avec plusieurs partitions : les requêtes qui génèrent un grand nombre de partitions dynamiques peuvent échouer et les requêtes qui effectuent des sélections sur des tables comportant un grand nombre de partitions peuvent nécessiter des temps d'exécution plus longs que prévu. Par exemple, une sélection effectuée sur 100 000 partitions peut prendre 10 minutes ou plus.

Fonctionnalités supplémentaires de Hive sur Amazon EMR

Amazon EMR étend Hive avec de nouvelles fonctionnalités qui prennent en charge l'intégration de Hive avec d'autres AWS services, tels que la capacité de lire et d'écrire sur Amazon Simple Storage Service (Amazon S3) et DynamoDB.

Variables dans Hive

Vous pouvez inclure des variables dans vos scripts en utilisant le symbole dollar et des accolades.

```
add jar ${LIB}/jsonserde.jar
```

Vous transmettez les valeurs de ces variables à Hive sur la ligne de commande à l'aide du paramètre `-d`, comme dans l'exemple suivant :

```
-d LIB=s3://elasticmapreduce/samples/hive-ads/lib
```

Vous pouvez également transmettre les valeurs via des étapes qui exécutent des scripts Hive.

Pour transmettre des valeurs de variables dans des étapes Hive à l'aide de la console

1. Ouvrez la console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com
2. Choisissez Créer un cluster.
3. Dans la section Étapes, pour Add Step (Ajouter étapes), choisissez Programme Hive dans la liste et Configurer et ajouter.
4. Dans la boîte de dialogue Ajouter une étape, spécifiez les paramètres en utilisant le tableau suivant comme guide, puis choisissez Ajouter.

Champ	Action
Emplacement S3 du script*	Spécifiez l'URI où votre script réside dans Amazon S3. La valeur doit être au format*BucketName* /*path*/*ScriptName* . Par exemple : `s3://elasticmapreduce/samples/hive-ads/libs/response-time-stats.q` .
Emplacement S3 d'entrée	Le cas échéant, spécifiez l'URI de l'emplacement de vos fichiers d'entrée dans Amazon S3. La valeur doit être sous la forme*BucketName* /*path*/. Si spécifié, elle sera transmise au script Hive comme un paramètre nommé INPUT. Par exemple : `s3://elasticmapreduce/samples/hive-ads/tables/` .
Emplacement S3 de sortie	Le cas échéant, spécifiez l'URI où vous voulez que le résultat soit stocké dans Amazon S3. La valeur doit être au format*BucketName* /*path*. Si spécifié, elle sera transmise au script Hive comme un paramètre nommé OUTPUT. Par exemple : `s3://amzn-s3-demo-bucket/hive-ads/output/` .
Arguments	Le cas échéant, entrez une liste d'arguments (chaînes séparées par des espaces) à transmettre à Hive. Si vous avez défini une variable de chemin d'accès dans votre script Hive nommé ${SAMPLE}, par exemple : `CREATE EXTERNAL TABLE logs (requestBeginTime STRING, requestEndTime STRING, hostname STRING) PARTITIONED BY (dt STRING) \ ROW FORMAT serde 'com.amazon.elasticmapreduce.JsonSerde' WITH SERDEPROPERTIES ( 'paths'='requestBeginTime, requestEndTime, hostname' ) LOCATION '${SAMPLE}/tables/impressions';` Pour transmettre une valeur pour la variable, saisissez ce qui suit dans la fenêtre Arguments : `-d SAMPLE=s3://elasticmapreduce/samples/hive-ads/` .

Champ	Action
Action sur échec	Cela détermine ce que le cluster fait en réponse à toute erreur. Les valeurs possibles pour ce paramètre sont : • Arrêter le cluster : si l'étape échoue, arrêtez le cluster. Si la protection de la résiliation du cluster est activée ET que la persistance est activée, il ne s'arrêtera pas. • Annuler et attendre : si l'étape échoue, annulez les étapes restantes. Si le cluster a la persistance activée, le cluster ne s'arrêtera pas. • Continuer : si l'étape échoue, passez à l'étape suivante.

5. Sélectionnez les valeurs si nécessaire, puis choisissez Créer un cluster.

Pour transmettre des valeurs variables aux étapes Hive à l'aide du AWS CLI

Pour transmettre des valeurs variables aux étapes Hive à l'aide du paramètre AWS CLI, utilisez le `--steps` paramètre et incluez une liste d'arguments.

-
> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=Hive Name=Pig --use-default-roles --ec2-attributes
 KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=Hive,Name="Hive Program",ActionOnFailure=CONTINUE,Args=[-f,s3://
elasticmapreduce/samples/hive-ads/libs/response-time-stats.q,-d,INPUT=s3://
elasticmapreduce/samples/hive-ads/tables,-d,OUTPUT=s3://amzn-s3-demo-bucket/hive-
ads/output/,-d,SAMPLE=s3://elasticmapreduce/samples/hive-ads/]
```

Pour plus d'informations sur l'utilisation des commandes Amazon EMR dans le AWS CLI, consultez. https://docs.aws.amazon.com/cli/latest/reference/emr

Pour transmettre des valeurs de variables dans des étapes Hive à l'aide de Java SDK

- L'exemple suivant montre comment transmettre des variables dans des étapes à l'aide du SDK. Pour plus d'informations, consultez la section Classe StepFactory dans la référence de AWS SDK pour Java l'API.

```
StepFactory stepFactory = new StepFactory();

   StepConfig runHive = new StepConfig()
     .withName("Run Hive Script")
     .withActionOnFailure("TERMINATE_JOB_FLOW")
     .withHadoopJarStep(stepFactory.newRunHiveScriptStep(“s3://amzn-s3-demo-bucket/
script.q”,
      Lists.newArrayList(“-d”,”LIB= s3://elasticmapreduce/samples/hive-ads/lib”));
```

Requêtes Amazon EMR Hive pour prendre en compte les schémas DynamoDB partiels

Amazon EMR Hive offre une flexibilité maximale lors de l'interrogation de tables DynamoDB en vous permettant de spécifier un sous-ensemble de colonnes sur lequel vous pouvez filtrer des données, plutôt que d'exiger que votre requête inclue toutes les colonnes. Cette technique de requête de schéma partiel est efficace lorsque vous avez un schéma de base de données fragmenté et que vous souhaitez filtrer des enregistrements en fonction de quelques colonnes, par exemple le filtrage des horodatages.

L'exemple suivant montre comment utiliser une requête Hive pour :

- Créez une table DynamoDB.
- Sélectionner un sous-ensemble d'éléments (lignes) dans DynamoDB et réduire les données à certaines colonnes.
- Copier les données qui en résultent dans Amazon S3.

```
DROP TABLE dynamodb;
DROP TABLE s3;

CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, fullColumn
 map<String, String>)
    STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
```

```
    TBLPROPERTIES (
     "dynamodb.table.name" = "myTable",
     "dynamodb.throughput.read.percent" = ".1000",
     "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");

CREATE EXTERNAL TABLE s3(map<String, String>)
     ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
     LOCATION 's3://bucketname/path/subpath/';

INSERT OVERWRITE TABLE s3 SELECT item fullColumn FROM dynamodb WHERE recordTimeStamp <
 "2012-01-01";
```

La table suivante illustre la syntaxe de requête pour la sélection de n'importe quelle combinaison d'éléments de DynamoDB.

Exemple de requête	Description du résultat
SÉLECTIONNEZ* PARMI *table_name* ;	Sélectionne tous les éléments (lignes) dans une table donnée et inclut des données à partir de toutes les colonnes disponibles pour ces éléments.
SÉLECTIONNEZ * D'*table_name* OÙ *field_name* = *value* ;	Sélectionne certains éléments (lignes) dans une table donnée et inclut des données à partir de toutes les colonnes disponibles pour ces éléments.
SÉLECTIONNEZ *column1_name column2_name* , *column3_name* PARMI *table_name* ;	Sélectionne tous les éléments (lignes) dans une table donnée et inclut des données à partir de certaines colonnes disponibles pour ces éléments.
SÉLECTIONNEZ *column1_name column2_name* , *column3_name* D'*table_name* OÙ *field_name* = *value* ;	Sélectionne certains éléments (lignes) dans une table donnée et inclut des données à partir de certaines colonnes disponibles pour ces éléments.

Copie des données entre les tables DynamoDB dans différentes régions AWS

Amazon EMR Hive fournit une propriété `dynamodb.region` que vous pouvez définir pour chaque table DynamoDB. Lorsque `dynamodb.region` est définie différemment sur deux tables, toute données que vous copiez entre les tables se produit automatiquement entre les régions spécifiées.

L'exemple suivant vous montre comment créer une table DynamoDB avec un script Hive qui définit la propriété `dynamodb.region` :

> **Note**
>
> Les propriétés de région par table remplacent les propriétés Hive globales.

```
CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, map<String,
 String> fullColumn)
    STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
    TBLPROPERTIES (
     "dynamodb.table.name" = "myTable",
     "dynamodb.region" = "eu-west-1",
     "dynamodb.throughput.read.percent" = ".1000",
     "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");
```

Définition des valeurs de débit de DynamoDB par table

Hive Amazon EMR vous permet de définir la DynamoDB readThroughputPercent et les writeThroughputPercent paramètres sur une base par table dans la définition de table. Le script Hive Amazon EMR suivant montre comment définir les valeurs de débit. Pour plus d'informations sur les valeurs de débit DynamoDB, consultez Spécification d'exigences de lecture et d'écriture pour des tables.

```
CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, map<String,
 String> fullColumn)
    STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
    TBLPROPERTIES (
     "dynamodb.table.name" = "myTable",
     "dynamodb.throughput.read.percent" = ".4",
     "dynamodb.throughput.write.percent" = "1.0",
     "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");
```

Configuration d'un metastore externe pour Hive

Par défaut, Hive enregistre les informations de metastore dans une base de données MySQL sur le système de fichiers du nœud primaire. Le metastore contient une description de la table et des données sous-jacentes sur lesquelles il est construit, notamment les noms de partition, les types de

données, etc. Lorsqu'un cluster est suspendu, tous les nœuds de cluster sont arrêtés, y compris le nœud primaire. Lorsque cela arrive, les données locales sont perdues, car les systèmes de fichiers du nœud utilisent un magasin éphémère. Pour que le metastore persiste, vous devez créer un metastore externe qui existe en dehors du cluster.

Vous avez deux options pour un metastore externe :

- AWS Le catalogue de données Glue (Amazon EMR version 5.8.0 ou ultérieure uniquement).

  Pour de plus amples informations, veuillez consulter Utilisation du catalogue de données AWS Glue comme métastore pour Hive.
- Amazon RDS ou Amazon Aurora.

  Pour de plus amples informations, veuillez consulter Utilisation d'une base de données MySQL externe ou d'Amazon Aurora.

> **Note**
>
> Si vous utilisez Hive 3 et que vous rencontrez trop de connexions au métastore Hive, configurez le paramètre `datanucleus.connectionPool.maxPoolSize` pour qu'il ait une valeur inférieure ou augmentez le nombre de connexions que le serveur de base de données peut gérer. L'augmentation du nombre de connexions est due à la façon dont Hive calcule le nombre maximum de connexions JDBC. Pour calculer la valeur optimale en termes de performances, consultez la section Propriétés de configuration de Hive.

Utilisation du catalogue de données AWS Glue comme métastore pour Hive

En utilisant les versions 5.8.0 ou ultérieures d'Amazon EMR, vous pouvez configurer Hive pour qu'il utilise le catalogue de données AWS Glue comme métastore. Cette configuration est recommandée si vous avez besoin d'un metastore persistant ou d'un metastore partagé par différents clusters, services, applications ou comptes AWS .

AWS Glue est un service d'extraction, transformation et chargement (ETL) entièrement géré qui facilite et rend abordable le classement par catégorie, le nettoyage, l'enrichissement et le déplacement des données de manière fiable entre différentes banques de données. Le catalogue de données AWS Glue fournit un référentiel de métadonnées unifié pour une variété de sources et de formats de données, s'intégrant à Amazon EMR ainsi qu'à Amazon RDS, Amazon Redshift

Spectrum, Athena, et toute application compatible avec le métastore Apache Hive. AWS Les crawlers Glue peuvent automatiquement déduire les schémas des données sources dans Amazon S3 et stocker les métadonnées associées dans le catalogue de données. Pour plus d'informations sur le catalogue de données, consultez Remplissage du catalogue de données AWS Glue dans le Guide du développeur AWS Glue.

Des frais distincts s'appliquent pour AWS Glue. Il y a un tarif mensuel pour le stockage et l'accès aux métadonnées dans le catalogue de données, un tarif horaire facturé à la minute pour les tâches ETL AWS Glue et l'exécution du analyseur, et un prix horaire facturé à la minute pour chaque point de terminaison de développement provisionné. Le catalogue de données vous permet de stocker jusqu'à un million d'objets sans aucuns frais. Au-delà d'un million d'objets, des frais de 1 USD vous seront facturés tous les 100 000 objets. Dans le catalogue de données, sont considérés comme des objets une table, une partition ou une base de données. Pour de plus amples informations, consultez Tarification Glue.

> **Important**
>
> Si vous avez créé des tables à l'aide d'Amazon Athena ou d'Amazon Redshift Spectrum avant le 14 août 2017, les bases de données et les tables sont stockées dans un catalogue géré par Athena, qui est distinct du catalogue de données Glue. AWS Pour intégrer Amazon EMR à ces tables, vous devez mettre à niveau vers le catalogue de données AWS Glue. Pour plus d'informations, consultez la section Mise à niveau vers le catalogue de données AWS Glue dans le guide de l'utilisateur Amazon Athena.

Spécifier le catalogue de données AWS Glue comme métastore

Vous pouvez spécifier le catalogue de données AWS Glue comme métastore à l'aide de la AWS Management Console AWS CLI, l'ou l'API Amazon EMR. Lorsque vous utilisez la CLI ou l'API, vous utilisez la classification de configuration de Hive pour spécifier le catalogue de données. En outre, avec Amazon EMR 5.16.0 et les versions ultérieures, vous pouvez utiliser la classification de configuration pour spécifier un catalogue de données dans un différent. Compte AWS Lorsque vous utilisez la console, vous pouvez spécifier le catalogue de données à l'aide des Options avancées ou des Options rapides.

Console

Pour spécifier le catalogue de données AWS Glue comme métastore Hive dans la console

1. Connectez-vous à la et ouvrez AWS Management Console la console Amazon EMR à https://console.aws.amazon.com l'adresse /emr.
2. Sous EMR activé EC2 dans le volet de navigation de gauche, choisissez Clusters, puis Create cluster.
3. Sous Ensemble d'applications, choisissez Core Hadoop ou HBaseCustom. Si vous personnalisez votre cluster, assurez-vous de sélectionner Hive ou l'une HCatalog de vos applications.
4. Sous Paramètres du catalogue de données AWS Glue, cochez la case Utiliser les métadonnées des tables Hive.
5. Choisissez toutes les autres options qui s'appliquent à votre cluster.
6. Pour lancer cluster, choisissez Créer un cluster.

CLI

Pour spécifier le catalogue de données AWS Glue comme métastore Hive avec l' AWS CLI

Pour plus d'informations sur la spécification d'une classification de configuration à l'aide de la AWS CLI et de l'API EMR, voir la rubrique. Configuration des applications

- Spécifiez la valeur de `hive.metastore.client.factory.class` à l'aide de la classification de configuration `hive-site` comme illustré dans l'exemple suivant :

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory"
    }
  }
]
```

Sur les versions EMR 5.28.0, 5.28.1, 5.29.0 ou 6.x, si vous créez un cluster utilisant le catalogue de données AWS Glue comme métastore, définissez la sur. `hive.metastore.schema.verification` `false` Cela empêche Hive et HCatalog de valider le schéma du métastore par rapport à MySQL. Sans cette configuration, le groupe d'instances principales sera suspendu après la reconfiguration sur Hive ou. HCatalog

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.schema.verification": "false"
    }
  }
]
```

Si vous disposez déjà d'un cluster sur EMR version 5.28.0, 5.28.1 ou 5.29.0, vous pouvez définir le groupe d'instances maître `hive.metastore.schema.verification` sur `false` avec les informations suivantes :

```
    Classification = hive-site
    Property       = hive.metastore.schema.verification
    Value          = false
```

Pour spécifier un catalogue de données dans un différent AWS compte, ajoutez la `hive.metastore.glue.catalogid` propriété, tel qu'illustré dans l'exemple suivant. Remplacez *`acct-id`* par le compte AWS du catalogue de données.

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.schema.verification": "false",
```

```
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

Autorisations IAM

Le profil d' EC2 instance pour un cluster doit avoir des autorisations IAM pour les actions AWS Glue. En outre, si vous activez le chiffrement pour les objets du catalogue de données AWS Glue, le rôle doit également être autorisé à chiffrer, déchiffrer et générer la AWS KMS key utilisée pour le chiffrement.

Autorisations pour les actions AWS Glue

Aucune action n'est requise si vous utilisez le profil d' EC2 instance par défaut pour Amazon EMR. La politique `AmazonElasticMapReduceforEC2Role` gérée attachée au `EMR_EC2_DefaultRole` autorise toutes les actions AWS Glue nécessaires. Toutefois, si vous spécifiez un profil d' EC2 instance et des autorisations personnalisés, vous devez configurer les actions AWS Glue appropriées. Utilisez la stratégie gérée `AmazonElasticMapReduceforEC2Role` comme point de départ. Pour plus d'informations, consultez la section Rôle de service pour les EC2 instances de cluster (profil d'EC2 instance) dans le guide de gestion Amazon EMR.

Autorisations de chiffrement et de déchiffrement du catalogue de données Glue AWS

Votre profil d'instance a besoin d'autorisations pour chiffrer et déchiffrer les données au moyen de votre clé. Vous n'avez pas besoin de configurer ces autorisations si les deux déclarations suivantes s'appliquent :

- Vous activez le chiffrement des objets du catalogue de données AWS Glue à l'aide des clés gérées pour AWS Glue.
- Vous utilisez un cluster identique au catalogue Compte AWS de données AWS Glue.

Sinon, vous devez ajouter l'instruction suivante à la politique d'autorisations attachée à votre profil d' EC2 instance.

```
[
    {
```

```
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "kms:Decrypt",
                    "kms:Encrypt",
                    "kms:GenerateDataKey"
                ],
                "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
            }
        ]
    }
]
```

Pour plus d'informations sur le chiffrement du catalogue de données AWS Glue, consultez Chiffrement de votre catalogue de données dans le Guide du développeur AWS Glue.

Autorisations basées sur les ressources

Si vous utilisez AWS Glue en conjonction avec Hive, Spark ou Presto dans Amazon EMR, AWS Glue prend en charge les politiques basées sur les ressources pour contrôler l'accès aux ressources du catalogue de données. Ces ressources comprennent les bases de données, les tables, les connexions et les fonctions définies par l'utilisateur. Pour plus d'informations, consultez Politiques de ressources AWS Glue dans le Guide du développeur AWS Glue.

Lorsque vous utilisez des politiques basées sur les ressources pour limiter l'accès à AWS Glue depuis Amazon EMR, le principal que vous spécifiez dans la politique d'autorisations doit être l'ARN de rôle associé au profil d' EC2 instance qui est spécifié lors de la création d'un cluster. Par exemple, pour une politique basée sur les ressources attachée à un catalogue, vous pouvez spécifier l'ARN de rôle pour le rôle de service par défaut pour les EC2 instances de cluster, *EMR_EC2_DefaultRole* en utilisant le format indiqué dans l'exemple suivant : `Principal`

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

L'ID de compte Glue *acct-id* peut être différent de l'ID de compte AWS Glue. Cela permet d'accéder aux clusters EMR à partir de comptes différents. Vous pouvez spécifier plusieurs principaux, chacun provenant d'un compte différent.

Considérations relatives à l'utilisation du catalogue de données AWS Glue

Tenez compte des éléments suivants lorsque vous utilisez le catalogue de données AWS Glue comme métastore avec Hive :

- L'ajout d'un auxiliaire JARs à l'aide du shell Hive n'est pas pris en charge. Pour contourner le problème, utilisez la classification `hive-site` de configuration pour définir la `hive.aux.jars.path` propriété, qui ajoute un auxiliaire JARs au chemin de classe Hive.
- Les transactions Hive ne sont pas prises en charge.
- Renommer des tables à partir d' AWS Glue n'est pas pris en charge.
- Lorsque vous créez une table Hive sans spécifier de `LOCATION`, les données de la table sont stockées à l'emplacement spécifié par la propriété `hive.metastore.warehouse.dir`. Par défaut, il s'agit d'un emplacement dans HDFS. Si un autre cluster doit accéder à la table, il échoue, sauf s'il dispose d'autorisations adéquates sur le cluster qui a créé la table. En outre, le stockage dans HDFS étant transitoire, si le cluster est résilié, les données de la table sont perdues et la table doit être recréée. Nous vous recommandons de spécifier un `LOCATION` dans Amazon S3 lorsque vous créez une table Hive à l'aide d' AWS Glue. Vous pouvez également utiliser la classification de configuration `hive-site` pour spécifier un emplacement dans Amazon S3 pour `hive.metastore.warehouse.dir`, qui s'applique à toutes les tables Hive. Si une table est créée dans un emplacement HDFS et que le cluster qui l'a créée est toujours en cours d'exécution, vous pouvez mettre à jour l'emplacement de la table vers Amazon S3 depuis AWS Glue. Pour plus d'informations, consultez Utilisation des tables dans la console AWS Glue dans le Guide du développeur AWS Glue.
- Les valeurs de partition contenant des guillemets et des apostrophes ne sont pas prises en charge, par exemple `PARTITION (owner="Doe's").`.
- Les statistiques sur les colonnes sont prises en charge à partir de la version emr-5.31.0.
- L'utilisation de l'autorisation Hive n'est pas prise en charge. Comme alternative, envisagez d'utiliser les politiques basées sur les ressources AWS Glue. Pour plus d'informations, consultez Utilisation de politiques basées sur les ressources pour l'accès d'Amazon EMR au catalogue de données Glue AWS.
- Les contraintes Hive ne sont pas prises en charge.
- L'optimisation basée sur les coûts dans Hive n'est pas prise en charge.
- Le paramètre `hive.metastore.partition.inherit.table.properties` n'est pas pris en charge.

- L'utilisation des constantes de métastore suivantes n'est pas prise en charge : `BUCKET_COUNT, BUCKET_FIELD_NAME, DDL_TIME, FIELD_TO_DIMENSION, FILE_INPUT_FORMAT, FILE_OUTPUT_FORMAT, HIVE_FILTER_FIELD_LAST_ACCESS, HIVE_FILTER_FIELD_OWNER, HIVE_FILTER_FIELD_PARAMS, IS_ARCHIVED, META_TABLE_COLUMNS, META_TABLE_COLUMN_TYPES, META_TABLE_DB, META_TABLE_LOCATION, META_TABLE_NAME, META_TABLE_PARTITION_COLUMNS, META_TABLE_SERDE, META_TABLE_STORAGE, ORIGINAL_LOCATION`.
- Lorsque vous utilisez une expression de prédicat, les valeurs explicites doivent se trouver sur le côté droit de l'opérateur de comparaison, sinon les requêtes risquent d'échouer.
  - Correct : `SELECT * FROM mytable WHERE time > 11`
  - Incorrect : `SELECT * FROM mytable WHERE 11 > time`
- Les versions 5.32.0 et 6.3.0 et ultérieures d'Amazon EMR prennent en charge l'utilisation de fonctions définies par l'utilisateur () dans les expressions de prédicat. UDFs Lorsque vous utilisez des versions antérieures, vos requêtes peuvent échouer en raison de la manière dont Hive essaie d'optimiser l'exécution des requêtes.
- Les tables temporaires ne sont pas prises en charge.
- Nous vous recommandons de créer des tableaux à l'aide d'applications via Amazon EMR plutôt que de les créer directement à l'aide d'Glue AWS . La création d'une table via AWS Glue peut entraîner l'absence de champs obligatoires et des exceptions de requête.
- Dans EMR 5.20.0 ou version ultérieure, l'élimination parallèle des partitions est activée automatiquement pour Spark et Hive lorsque AWS Glue Data Catalog est utilisé comme métastore. Cette modification réduit considérablement le temps de planification des requêtes en exécutant plusieurs requêtes en parallèle pour récupérer des partitions. Le nombre total de segments pouvant être exécutés simultanément est compris entre 1 et 10. La valeur par défaut est 5, ce qui est recommandé. Vous pouvez le modifier en spécifiant la propriété `aws.glue.partition.num.segments` dans la classification de configuration `hive-site`. En cas de limitation, vous pouvez désactiver la fonctionnalité en remplaçant la valeur par 1. Pour en savoir plus, consultez Structure d'un segment AWS Glue.

Utilisation d'une base de données MySQL externe ou d'Amazon Aurora

Pour utiliser une base de données externe MySQL ou Amazon Aurora en tant que metastore Hive, vous devez substituer les valeurs de configuration par défaut pour le metastore dans Hive pour spécifier l'emplacement de la base de données externe sur une instance MySQL Amazon RDS ou une Amazon RDS ou une Amazon Aurora Aurora Aurora Aurora Aurora Aurora Aurora Aurora Aurora

Aurora Aurora Aurora

ou ceux dotés d'informations d'identification déléguées) l'accès aux services sur ce cluster en créant un rôle qui refuse explicitement les autorisations pour la clé d'API `elasticmapreduce:DescribeCluster`.

a. Créez un fichier de configuration appelé `hiveConfiguration.json` contenant les modifications apportées à `hive-site.xml` comme illustré dans l'exemple suivant.

   Remplacez *`hostname`* par l'adresse DNS de votre instance Amazon RDS exécutant la base de données, *`username`* et *`password`* par les informations d'identification de votre base de données. Pour plus d'informations sur la connexion aux instances de base de données MySQL et Aurora, consultez Connexion à une instance de base de données exécutant le moteur de base de données MySQL et Connexion à un cluster de base de données Athena dans le Guide de l'utilisateur Amazon RDS. `javax.jdo.option.ConnectionURL` est la chaîne de connexion JDBC d'un metastore JDBC. `javax.jdo.option.ConnectionDriverName` est le nom de classe de pilote d'un metastore JDBC.

   Les pilotes JDBC de MySQL sont installés par Amazon EMR.

   La propriété value ne peut pas contenir d'espaces ni de retours chariot. Elle doit figurer entièrement sur une seule ligne.

   ```
   [
       {
         "Classification": "hive-site",
         "Properties": {
           "javax.jdo.option.ConnectionURL": "jdbc:mysql://hostname:3306/hive?
   createDatabaseIfNotExist=true",
           "javax.jdo.option.ConnectionDriverName": "org.mariadb.jdbc.Driver",
           "javax.jdo.option.ConnectionUserName": "username",
           "javax.jdo.option.ConnectionPassword": "password"
         }
       }
     ]
   ```

b. Référencez le `hiveConfiguration.json` fichier lorsque vous créez le cluster, comme indiqué dans la AWS CLI commande suivante. Le fichier est stocké localement dans cette commande. Vous pouvez également charger le fichier sur Amazon S3

et le référencer depuis cet endroit, par exemple, `s3://DOC-EXAMPLE-BUCKET/hiveConfiguration.json`.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=Hive --configurations file://hiveConfiguration.json --use-
default-roles
```

4. Connectez-vous au nœud primaire de votre cluster.

   Pour plus d'informations sur la façon de se connecter au nœud primaire, consultez Connexion au nœud primaire à l'aide de SSH dans le Guide de gestion d'Amazon EMR.

5. Créez vos tables Hive en spécifiant l'emplacement sur Amazon S3 en entrant une commande similaire à :

```
CREATE EXTERNAL TABLE IF NOT EXISTS table_name
(
key int,
value int
)
LOCATION s3://DOC-EXAMPLE-BUCKET/hdfs/
```

6. Ajoutez votre script Hive au cluster en cours d'exécution.

Votre cluster Hive s'exécute en utilisant le métastore situé dans Amazon RDS. Lancez tous les clusters Hive supplémentaires qui partagent ce metastore en spécifiant l'emplacement du metastore.

Utilisation du pilote Hive JDBC

Vous pouvez utiliser des outils de business intelligence populaires tels que Microsoft Excel et Tableau avec Amazon EMR pour explorer et visualiser vos données. MicroStrategy QlikView Un

grand nombre de ces outils ont besoin d'un pilote JDBC (Java DataBase Connectivity) ou ODBC (Open DataBase Connectivity). Amazon EMR prend en charge la connectivité JDBC et ODBC.

L'exemple ci-dessous explique comment utiliser SQL Workbench/J en tant que client SQL pour vous connecter à un cluster Hive dans Amazon EMR. Pour obtenir des pilotes supplémentaires, consultez Utilisation des outils d'aide à la décision avec Amazon EMR.

Avant d'installer et d'utiliser SQL Workbench/J, téléchargez le package du pilote et installez ce dernier. Les pilotes inclus dans le package prennent en charge les versions Hive disponibles dans Amazon EMR versions 4.0 et versions ultérieures. Pour obtenir des notes de mise à jour et une documentation détaillées, consultez la documentation PDF incluse dans le package.

- Téléchargement du dernier package de pilotes Hive JDBC

  http://awssupportdatasvcs.com/bootstrap-actions/Simba/latest/
- Anciennes versions du pilote JDBC Hive

  http://awssupportdatasvcs.com/bootstrap-actions/Simba/

Pour installer et configurer SQL Workbench

1. Téléchargez le client SQL Workbench/J pour votre système d'exploitation à l'adresse http://www.sql-workbench.net/downloads.html.
2. Installez SQL Workbench/J. Pour plus d'informations, consultez Installing and starting SQL Workbench/J dans le manuel de l'utilisateur SQL Workbench/J.
3. Utilisateurs Linux, Unix et Mac OS X : dans une session de terminal, créez un tunnel SSH vers le nœud principal de votre cluster à l'aide de la commande suivante. Remplacez-le *master-public-dns-name* par le nom DNS public du nœud principal et *path-to-key-file* par l'emplacement et le nom de fichier de votre fichier de clé EC2 privée Amazon (`.pem`).

   ```
   ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10000:localhost:10000
    hadoop@master-public-dns-name
   ```

   Utilisateurs Windows : Dans une session PuTTY, créez un tunnel SSH pour le nœud principal de votre cluster (à l'aide du réacheminement de port local) avec `10000` pour Port source et *master-public-dns-name*`:10000` pour Destination. Remplacez *master-public-dns-name* par le nom DNS public du nœud maître.
4. Ajoutez le pilote JDBC à SQL Workbench.

a. Dans la boîte de dialogue Select Connection Profile (Sélectionner le profil de connexion), cliquez sur Manage Drivers (Gérer les pilotes).

b. Cliquez sur l'icône Create a new entry (Créer une nouvelle entrée) (page blanche).

c. Dans le champ Name (Nom), saisissez **`Hive JDBC`**.

d. Pour Library (Bibliothèque), cliquez sur l'icône Select the JAR file(s) (Sélectionner le(s) fichier(s) JAR).

e. Accédez à l'emplacement contenant les pilotes extraits. Sélectionnez les pilotes inclus dans la version du package de pilotes JDBC que vous avez téléchargée, puis cliquez sur Ouvrir.

   Par exemple, votre package de pilotes JDBC peut inclure les éléments suivants. JARs

   ```
   hive_metastore.jar
   hive_service.jar
   HiveJDBC41.jar
   libfb303-0.9.0.jar
   libthrift-0.9.0.jar
   log4j-1.2.14.jar
   ql.jar
   slf4j-api-1.5.11.jar
   slf4j-log4j12-1.5.11.jar
   TCLIServiceClient.jar
   zookeeper-3.4.6.jar
   ```

f. Dans la boîte de dialogue Please select one driver (Veuillez sélectionner un pilote), sélectionnez `com.amazon.hive.jdbc41.HS2Driver`, OK.

5. Lorsque vous revenez à la boîte de dialogue Manage Drivers (Gérer les pilotes) vérifiez que le champ Classname (Nom de classe) est renseigné et cliquez sur OK.

6. Lorsque vous revenez à la boîte de dialogue Select Connection Profile (Sélectionner le profil de connexion), vérifiez que le champ Driver (Pilote) est défini sur Hive JDBC et fournissez la chaîne de connexion JDBC suivante dans le champ URL : `jdbc:hive2://localhost:10000/default`.

7. Sélectionnez OK pour vous connecter. Une fois la connexion établie, les détails de connexion apparaissent en haut de la fenêtre SQL Workbench/J.

Pour plus d'informations sur l'utilisation de Hive et de l'interface JDBC, consultez HiveClientet Hive JDBCInterface dans la documentation d'Apache Hive.

Amélioration des performances de Hive

Amazon EMR offre des fonctionnalités permettant d'optimiser les performances lors de l'utilisation de Hive pour interroger, lire et écrire des données sauvegardées dans Amazon S3.

S3 Select peut améliorer la performance de requête pour les fichiers CSV et JSON dans certaines applications en « poussant vers le bas » le traitement à Amazon S3.

Le validateur optimisé EMRFS S3 est une alternative à la OutputCommitterclasse, qui élimine les opérations de liste et de renommage afin d'améliorer les performances lors de l'écriture de fichiers Amazon S3 à l'aide d'EMRFS.

Rubriques

- Activation du validateur optimisé Hive EMRFS S3
- Utilisation de S3 Select avec Hive pour améliorer les performances
- Optimisation MSCK

Activation du validateur optimisé Hive EMRFS S3

Le Hive EMRFS S3 Optimized Committer est une méthode alternative permettant à EMR Hive d'écrire des fichiers pour des requêtes d'insertion lors de l'utilisation d'EMRFS. Le Committer élimine les opérations de liste et de renommage effectuées sur Amazon S3 et améliore les performances de l'application. La fonctionnalité est disponible à partir de EMR 5.34 et EMR 6.5.

Activation du contributeur

Si vous souhaitez activer EMR Hive pour utiliser `HiveEMRFSOptimizedCommitter` pour valider les données par défaut pour toutes les tables gérées et externes par Hive, utilisez la configuration `hive-site` suivante dans les clusters EMR 6.5.0 ou EMR 5.34.0.

```
[
   {
      "classification": "hive-site",
      "properties": {
         "hive.blobstore.use.output-committer": "true"
      }
   }
]
```

> **Note**
>
> N'activez pas cette fonction lorsque `hive.exec.parallel` est défini sur `true`.

Limites

Les restrictions de base suivantes s'appliquent aux balises :

- L'activation de Hive pour fusionner automatiquement de petits fichiers n'est pas prise en charge. La logique de validation Hive par défaut sera utilisée même lorsque le validateur optimisé est activé.
- Les tables Hive ACID ne sont pas prises en charge. La logique de validation Hive par défaut sera utilisée même lorsque le validateur optimisé est activé.
- La nomenclature des noms de fichiers pour les fichiers écrits est passée de Hive `<task_id>_<attempt_id>_<copy_n>` à `<task_id>_<attempt_id>_<copy_n>_<query_id>`. Par exemple, un fichier nommé `s3://warehouse/table/partition=1/000000_0` sera remplacé par `s3://warehouse/table/partition=1/000000_0-hadoop_20210714130459_ba7c23ec-5695-4947-9d98-8a40ef759222-1`. Le `query_id` ici est une combinaison du nom d'utilisateur, de l'horodatage et de l'UUID.
- Lorsque des partitions personnalisées se trouvent sur des systèmes de fichiers différents (HDFS, S3), cette fonctionnalité est automatiquement désactivée. La logique de validation Hive par défaut sera utilisée lorsqu'elle sera activée.

Utilisation de S3 Select avec Hive pour améliorer les performances

> **Important**
>
> Amazon S3 Select n'est plus disponible pour les nouveaux clients. Les clients existants d'Amazon S3 Select peuvent continuer à utiliser cette fonctionnalité comme d'habitude. En savoir plus

À partir de la version 5.18.0 d'Amazon EMR, vous pouvez utiliser S3 Select avec Hive sur Amazon EMR. S3 Select permet aux applications de ne récupérer qu'un sous-ensemble de données d'un objet. Pour Amazon EMR, le travail de calcul impliqué dans le filtrage d'ensembles larges de données

pour le traitement est « poussé en bas », du cluster vers Amazon S3, ce qui peut améliorer la performance de certaines applications et peut réduire la quantité de données transférée entre Amazon EMR et Amazon S3.

S3 Select est pris en charge avec les tables Hive en fonction des fichiers JSON et CSV, et en définissant la variable de configuration `s3select.filter` sur `true` durant votre session Hive. Pour plus d'informations et d'exemples, consultez Spécification de S3 Select dans votre code.

S3 Select est-il adapté à mon application ?

Nous vous recommandons de tester vos applications avec et sans S3 Select pour voir si son utilisation peut être adaptée à votre application.

Utilisez les consignes suivantes pour déterminer si votre application peut utiliser S3 Select :

- Votre requête exclut plus de la moitié de l'ensemble de données d'origine.
- Vos prédicats de filtre de requête utilisent les colonnes qui ont un type de données pris en charge par Amazon S3 Select. Pour plus d'informations, consultez Types de données dans le Guide de l'utilisateur Amazon Simple Storage Service.
- Votre connexion réseau entre Amazon S3 et le cluster Amazon EMR a une bonne vitesse de transfert et une bande passante disponible. Amazon S3 ne compresse pas les réponses HTTP, de sorte que la taille de la réponse est susceptible d'augmenter pour les fichiers d'entrée compressés.

Considérations et restrictions

- Le chiffrement côté serveur d'Amazon S3 avec des clés de chiffrement fournies par le client (SSE-C) et le chiffrement côté client ne sont pas pris en charge.
- La propriété `AllowQuotedRecordDelimiters` n'est pas prise en charge. Si cette propriété est spécifiée, la requête échoue.
- Seuls les fichiers CSV et JSON au format UTF-8 sont pris en charge. Le format multiligne CSVs et le format JSON ne sont pas pris en charge.
- Seuls les fichiers non compressés ou gzip ou bzip2 sont pris en charge.
- Les caractères de commentaire ne sont pas pris en charge dans la dernière ligne.
- Des lignes vides à la fin d'un fichier ne sont pas traitées.
- Hive sur Amazon EMR prend en charge les types de données primitifs que S3 Select prend en charge. Pour plus d'informations, consultez Types de données dans le Guide de l'utilisateur Amazon Simple Storage Service.

Spécification de S3 Select dans votre code

Pour utiliser S3 Select dans votre table Hive, créez la table en spécifiant `com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat` comme nom de classe `INPUTFORMAT` et spécifiez une valeur pour la propriété `s3select.format` à l'aide de la clause `TBLPROPERTIES`.

S3 Select est désactivé par défaut lorsque vous exécutez des requêtes. Activez S3 Select en définissant `s3select.filter` sur `true` dans votre session Hive, comme illustré ci-dessous. Les exemples suivants montrent comment spécifier S3 Select lors de la création d'une table à partir de fichiers CSV et JSON sous-jacents, puis comment interroger la table à l'aide d'une instruction de sélection simple.

Example Instruction CREATE TABLE pour une table basée sur CSV

```
CREATE TABLE mys3selecttable (
col1 string,
col2 int,
col3 boolean
)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
STORED AS
INPUTFORMAT
  'com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat'
OUTPUTFORMAT
  'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'
LOCATION 's3://path/to/mycsvfile/'
TBLPROPERTIES (
  "s3select.format" = "csv",
  "s3select.headerInfo" = "ignore"
);
```

Example Instruction CREATE TABLE pour une table basée sur JSON

```
CREATE TABLE mys3selecttable (
col1 string,
col2 int,
col3 boolean
)
ROW FORMAT SERDE 'org.apache.hive.hcatalog.data.JsonSerDe'
STORED AS
```

```
INPUTFORMAT
  'com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat'
OUTPUTFORMAT
  'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'
LOCATION 's3://path/to/json/'
TBLPROPERTIES (
  "s3select.format" = "json"
);
```

Example Instruction SELECT TABLE

```
SET s3select.filter=true;
SELECT * FROM mys3selecttable WHERE col2 > 10;
```

Optimisation MSCK

Hive stocke une liste de partitions pour chaque table dans son métastore. Toutefois, lorsque des partitions sont directement ajoutées ou supprimées du système de fichiers, le métastore Hive n'est pas au courant de ces modifications. La commande MSCK met à jour les métadonnées de partition dans le métastore Hive pour les partitions directement ajoutées ou supprimées du système de fichiers. La syntaxe de cette commande est la suivante :

```
MSCK [REPAIR] TABLE table_name [ADD/DROP/SYNC PARTITIONS];
```

Hive implémente cette commande comme suit :

1. Hive récupère toutes les partitions de la table dans le métastore. À partir de la liste des chemins de partition qui n'existent pas dans le système de fichiers, crée une liste de partitions à supprimer du métastore.
2. Hive rassemble les chemins de partition présents dans le système de fichiers, les compare à la liste des partitions du métastore et génère une liste des partitions qui doivent être ajoutées au métastore.
3. Hive met à jour le métastore en utilisant le mode `ADD`, `DROP` ou `SYNC`.

> **Note**
>
> Lorsque le métastore contient de nombreuses partitions, l'étape visant à vérifier si une partition n'existe pas dans le système de fichiers prend du temps, car l'appel d'API `exists` du système de fichiers doit être effectué pour chaque partition.

Dans Amazon EMR 6.5.0, Hive a introduit un indicateur appelé `hive.emr.optimize.msck.fs.check`. Lorsqu'il est activé, cet indicateur permet à Hive de vérifier la présence d'une partition dans la liste des chemins de partition du système de fichiers générée à l'étape 2 ci-dessus au lieu d'effectuer des appels d'API du système de fichiers. Dans Amazon EMR 6.8.0, Hive a activé cette optimisation par défaut, éliminant ainsi le besoin de définir l'indicateur `hive.emr.optimize.msck.fs.check`.

Utilisation de Live Long and Process (LLAP) de Hive

Amazon EMR 6.0.0 prend en charge la fonctionnalité Live Long and Process (LLAP) pour Hive. LLAP utilise des démons persistants avec une mise en cache intelligente en mémoire pour améliorer les performances des requêtes par rapport au mode d'exécution précédent du conteneur Tez par défaut.

Les démons Hive LLAP sont gérés et exécutés en tant que service YARN. Étant donné qu'un service YARN peut être considéré comme une application YARN longue durée, certaines de vos ressources de cluster sont dédiées à Hive LLAP et ne peuvent pas être utilisées pour d'autres charges de travail. Pour plus d'informations, consultez LLAP et API YARN Service.

Activer Hive LLAP sur Amazon EMR

Pour activer Hive LLAP sur Amazon EMR, fournissez la configuration suivante lorsque vous lancez un cluster.

```
[
  {
    "Classification": "hive",
    "Properties": {
      "hive.llap.enabled": "true"
    }
  }
]
```

Pour plus d'informations, consultez Configuration des applications.

Par défaut, Amazon EMR alloue environ 60 % des ressources YARN de cluster aux démons Hive LLAP. Vous pouvez configurer le pourcentage de ressource YARN de cluster allouée à Hive LLAP et le nombre de tâches et de nœuds principaux à prendre en compte pour l'allocation LLAP Hive.

Par exemple, la configuration suivante démarre Hive LLAP avec trois démons sur trois nœuds de tâches ou nœuds principaux et alloue 40 % de la ressource YARN de ces trois nœuds aux démons LLAP Hive.

```
[
  {
    "Classification": "hive",
    "Properties": {
      "hive.llap.enabled": "true",
      "hive.llap.percent-allocation": "0.4",
      "hive.llap.num-instances": "3"
    }
  }
]
```

Vous pouvez utiliser les configurations `hive-site` suivantes dans l'API de classification pour remplacer les paramètres de ressource LLAP par défaut.

Propriété	Description
hive.llap.daemon.yarn.container.mb	Taille totale du conteneur du démon LLAP (en Mo)
hive.llap.daemon.memory.per.instance.mb	Mémoire totale utilisée par les exécuteurs dans le conteneur de démon LLAP (en Mo)
hive.llap.io.memory.size	Taille du cache pour l'entrée ou la sortie LLAP
hive.llap.daemon.num.executors	Nombre d'exécuteurs par démon LLAP

Démarrez manuellement le LLAP de Hive sur votre cluster.

Toutes les dépendances et configurations utilisées par LLAP sont compressées dans l'archive tar LLAP dans le cadre du démarrage du cluster. Si LLAP est activé avec `"hive.llap.enabled": "true"`, nous vous recommandons d'utiliser la reconfiguration Amazon EMR pour apporter des modifications à la configuration LLAP.

Sinon, pour toute modification manuelle de `hive-site.xml`, vous devez reconstruire l'archive tar LLAP à l'aide de la commande `hive --service llap`, comme le montre l'exemple suivant.

```
# Define how many resources you want to allocate to Hive LLAP

LLAP_INSTANCES=<how many llap daemons to run on cluster>
LLAP_SIZE=<total container size per llap daemon>
LLAP_EXECUTORS=<number of executors per daemon>
LLAP_XMX=<Memory used by executors>
LLAP_CACHE=<Max cache size for IO allocator>

yarn app -enableFastLaunch

hive --service llap \
--instances $LLAP_INSTANCES \
--size ${LLAP_SIZE}m \
--executors $LLAP_EXECUTORS \
--xmx ${LLAP_XMX}m \
--cache ${LLAP_CACHE}m \
--name llap0 \
--auxhbase=false \
--startImmediately
```

Contrôle de statut de LLAP de Hive

Utilisez la commande suivante pour vérifier l'état de Hive LLAP via Hive.

```
hive --service llapstatus
```

Utilisez la commande suivante pour vérifier l'état de Hive LLAP à l'aide de YARN.

```
yarn app -status (name-of-llap-service)

# example:
```

```
yarn app -status llap0 | jq
```

Démarrage ou arrêt de LLAP de Hive

Étant donné que le LLAP de Hive s'exécute en tant que service YARN permanent, vous devez arrêter ou redémarrer le service YARN pour arrêter ou redémarrer le LLAP de Hive. Les commandes suivantes illustrent ce problème.

```
yarn app -stop llap0
yarn app -start llap0
```

Redimensionner le nombre de démons LLAP Hive

Utilisez la commande suivante pour réduire le nombre d'instances LLAP.

```
yarn app -flex llap0 -component llap -1
```

Pour plus d'informations, consultez Ajustement du composant d'un service.

Chiffrement dans Hive

Cette section décrit les types de chiffrement pris en charge par Amazon EMR.

Chiffrement modulaire du parquet dans Hive

Le chiffrement modulaire Parquet fournit un contrôle d'accès et un chiffrement au niveau des colonnes afin d'améliorer la confidentialité et l'intégrité des données stockées au format de fichier Parquet. Cette fonctionnalité est disponible dans Amazon EMR Hive à partir de la version 6.6.0.

Les solutions de sécurité et d'intégrité précédemment prises en charge, notamment le chiffrement de fichiers ou le chiffrement de la couche de stockage, sont décrites dans la section Options de chiffrement du guide de gestion Amazon EMR. Ces solutions peuvent être utilisées pour les fichiers Parquet, mais l'exploitation des nouvelles fonctionnalités du mécanisme de cryptage intégré de Parquet permet un accès granulaire au niveau des colonnes, ainsi que des améliorations en termes de performances et de sécurité. Pour en savoir plus sur cette fonctionnalité, consultez la page github d'Apache intitulée Parquet Modular Encryption.

Les utilisateurs transmettent les configurations aux lecteurs et rédacteurs Parquet à l'aide des configurations Hadoop. Les configurations détaillées permettant aux utilisateurs de configurer les

lecteurs et les enregistreurs pour activer le chiffrement ainsi que les fonctionnalités avancées sont documentées dans PARQUET-1854 : Properties-driven Interface to Parquet Encryption Management (Interface pilotée par les propriétés pour la gestion du chiffrement de Parquet).

Exemples d'utilisation

L'exemple suivant couvre la création et l'écriture dans une table Hive en utilisant AWS KMS pour gérer les clés de chiffrement.

1. Implémentez un KmsClient pour le AWS KMS service tel que décrit dans le document PARQUET-1373 : Outils de gestion des clés de chiffrement. L'exemple suivant montre un extrait de code d'implémentation.

```
package org.apache.parquet.crypto.keytools;

import com.amazonaws.AmazonClientException;
import com.amazonaws.AmazonServiceException;
import com.amazonaws.regions.Regions;
import com.amazonaws.services.kms.AWSKMS;
import com.amazonaws.services.kms.AWSKMSClientBuilder;
import com.amazonaws.services.kms.model.DecryptRequest;
import com.amazonaws.services.kms.model.EncryptRequest;
import com.amazonaws.util.Base64;
import org.apache.hadoop.conf.Configuration;
import org.apache.parquet.crypto.KeyAccessDeniedException;
import org.apache.parquet.crypto.ParquetCryptoRuntimeException;
import org.apache.parquet.crypto.keytools.KmsClient;
import org.slf4j.Logger;
import org.slf4j.LoggerFactory;

import java.nio.ByteBuffer;
import java.nio.charset.Charset;
import java.nio.charset.StandardCharsets;

public class AwsKmsClient implements KmsClient {

    private static final AWSKMS AWSKMS_CLIENT = AWSKMSClientBuilder
            .standard()
            .withRegion(Regions.US_WEST_2)
            .build();
    public static final Logger LOG = LoggerFactory.getLogger(AwsKmsClient.class);

    private String kmsToken;
```

```
   private Configuration hadoopConfiguration;

   @Override
   public void initialize(Configuration configuration, String kmsInstanceID, String
kmsInstanceURL, String accessToken) throws KeyAccessDeniedException {
       hadoopConfiguration = configuration;
       kmsToken = accessToken;

   }

   @Override
   public String wrapKey(byte[] keyBytes, String masterKeyIdentifier) throws
KeyAccessDeniedException {
       String value = null;
       try {
           ByteBuffer plaintext = ByteBuffer.wrap(keyBytes);

           EncryptRequest req = new
EncryptRequest().withKeyId(masterKeyIdentifier).withPlaintext(plaintext);
           ByteBuffer ciphertext = AWSKMS_CLIENT.encrypt(req).getCiphertextBlob();

           byte[] base64EncodedValue = Base64.encode(ciphertext.array());
           value = new String(base64EncodedValue, Charset.forName("UTF-8"));
       } catch (AmazonClientException ae) {
           throw new KeyAccessDeniedException(ae.getMessage());
       }
       return value;
   }

   @Override
   public byte[] unwrapKey(String wrappedKey, String masterKeyIdentifier) throws
KeyAccessDeniedException {
       byte[] arr = null;
       try {
           ByteBuffer ciphertext  =
ByteBuffer.wrap(Base64.decode(wrappedKey.getBytes(StandardCharsets.UTF_8)));
           DecryptRequest request = new
DecryptRequest().withKeyId(masterKeyIdentifier).withCiphertextBlob(ciphertext);
           ByteBuffer decipheredtext =
AWSKMS_CLIENT.decrypt(request).getPlaintext();
           arr = new byte[decipheredtext.remaining()];
           decipheredtext.get(arr);
       } catch (AmazonClientException ae) {
           throw new KeyAccessDeniedException(ae.getMessage());
```

```
        }
        return arr;
    }
}
```

2. Créez vos clés de AWS KMS chiffrement pour le pied de page ainsi que pour les colonnes auxquelles vos rôles IAM ont accès, comme décrit dans la section Création de clés dans le guide du AWS Key Management Service développeur. Le rôle IAM par défaut est EMR_ECS_Default.
3. Dans l'application Hive sur un cluster Amazon EMR, ajoutez le client ci-dessus à l'aide de l'instruction `ADD JAR`, comme décrit dans la documentation Apache Hive Resources. Voici un exemple de déclaration :

   ```
   ADD JAR 's3://location-to-custom-jar';
   ```

   Une autre méthode consiste à ajouter le fichier JAR à l'`auxlib` de Hive à l'aide d'une action d'amorçage. Voici un exemple de ligne à ajouter à l'action boostrap :

   ```
   aws s3 cp 's3://location-to-custom-jar' /usr/lib/hive/auxlib
   ```

4. Définissez les configurations suivantes :

   ```
   set
    parquet.crypto.factory.class=org.apache.parquet.crypto.keytools.PropertiesDrivenCryptoFacto
   set
    parquet.encryption.kms.client.class=org.apache.parquet.crypto.keytools.AwsKmsClient;
   ```

5. Créez une table Hive au format Parquet, spécifiez les AWS KMS clés dans SERDEPROPERTIES et insérez-y des données :

   ```
   CREATE TABLE my_table(name STRING, credit_card STRING)
   ROW FORMAT SERDE 'org.apache.hadoop.hive.ql.io.parquet.serde.ParquetHiveSerDe’
   WITH SERDEPROPERTIES (
     'parquet.encryption.column.key’=<aws-kms-key-id-for-column-1>: credit_card’,
     'parquet.encryption.footer.key’='<aws-kms-key-id-for-footer>’)
   STORED AS parquet
   LOCATION “s3://<bucket/<warehouse-location>/my_table”;

   INSERT INTO my_table SELECT
   java_method ('org.apache.commons.lang.RandomStringUtils','randomAlphabetic',5) as
    name,
   ```

```
java_method ('org.apache.commons.lang.RandomStringUtils','randomAlphabetic',10) as
 credit_card
from (select 1) x lateral view posexplode(split(space(100),' ')) pe as i,x;

select * from my_table;
```

6. Vérifiez que lorsque vous créez une table externe au même endroit sans accès aux AWS KMS clés (par exemple, l'accès au rôle IAM est refusé), vous ne pouvez pas lire les données.

```
CREATE EXTERNAL TABLE ext_table (name STRING, credit_card STRING)
ROW FORMAT SERDE 'org.apache.hadoop.hive.ql.io.parquet.serde.ParquetHiveSerDe’
STORED AS parquet
LOCATION “s3://<bucket>/<warehouse-location>/my_table”;

SELECT * FROM ext_table;
```

7. La dernière instruction doit générer l'exception suivante :

```
Failed with exception
 java.io.IOException:org.apache.parquet.crypto.KeyAccessDeniedException: Footer key:
 access denied
```

Chiffrement en transit en 2 HiveServer

À partir de la version 6.9.0 d'Amazon EMR, HiveServer 2 (HS2) est compatible TLS/SSL dans le cadre de la configuration de sécurité. Chiffrement en transit en 2 HiveServer Cela affecte la façon dont vous vous connectez à l' HS2 exécution sur un cluster Amazon EMR avec le chiffrement en transit activé. Pour vous connecter à HS2, vous devez modifier les valeurs `TRUSTSTORE_PATH` des `TRUSTSTORE_PASSWORD` paramètres et dans l'URL JDBC. L'URL suivante est un exemple de connexion JDBC pour HS2 avec les paramètres requis :

```
jdbc:hive2://HOST_NAME:10000/
default;ssl=true;sslTrustStore=TRUSTSTORE_PATH;trustStorePassword=TRUSTSTORE_PASSWORD
```

Utilisez les instructions appropriées pour le chiffrement sur cluster 2 ou hors cluster 2 ci-dessous HiveServer.

On-cluster HS2 access

Si vous accédez à HiveServer 2 à l'aide du client Beeline après avoir accédé au nœud principal en SSH, faites référence `/etc/hadoop/conf/ssl-server.xml` pour trouver les valeurs des `TRUSTSTORE_PASSWORD` paramètres `TRUSTSTORE_PATH` et à l'aide de la configuration `ssl.server.truststore.location` et. `ssl.server.truststore.password`

Les exemples de commandes suivants peuvent vous aider à récupérer ces configurations :

```
TRUSTSTORE_PATH=$(sed -n '/ssl.server.truststore.location/,+2p' /etc/hadoop/conf/
ssl-server.xml | awk -F "[><]" '/value/{print $3}')
TRUSTSTORE_PASSWORD=$(sed -n '/ssl.server.truststore.password/,+2p' /etc/hadoop/
conf/ssl-server.xml | awk -F "[><]" '/value/{print $3}')
```

Off-cluster HS2 access

Si vous accédez à HiveServer 2 depuis un client extérieur au cluster Amazon EMR, vous pouvez utiliser l'une des approches suivantes pour obtenir le et : `TRUSTSTORE_PATH` `TRUSTSTORE_PASSWORD`

- Convertissez le fichier PEM créé lors de la configuration de sécurité en fichier JKS et utilisez le même dans l'URL de connexion JDBC. Par exemple, avec openssl et keytool, utilisez les commandes suivantes :

  ```
  openssl pkcs12 -export -in trustedCertificates.pem -inkey privateKey.pem -out
   trustedCertificates.p12 -name "certificate"
  keytool -importkeystore -srckeystore trustedCertificates.p12 -srcstoretype pkcs12
   -destkeystore trustedCertificates.jks
  ```

- Vous pouvez également faire référence à `/etc/hadoop/conf/ssl-server.xml` pour trouver les valeurs des paramètres `TRUSTSTORE_PATH` et `TRUSTSTORE_PASSWORD` à l'aide de la configuration `ssl.server.truststore.location` et `ssl.server.truststore.password`. Téléchargez le fichier magasin de confiance sur la machine cliente et utilisez le chemin sur la machine cliente comme `TRUSTSTORE_PATH`.

  Pour plus d'informations sur l'accès aux applications depuis un client extérieur au cluster Amazon EMR, consultez Utiliser le pilote JDBC Hive.

Historique des versions de Hive

Le tableau suivant répertorie la version de Hive incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version de Hive

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-7.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7,8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
		resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7,7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-7.6.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.5.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-7.4.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.3.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-7.2.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-5.36.2	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-7.1.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.0.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.15.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.14.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.13.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.12.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.11.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.11.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.10.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.10.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.9.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.8.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-5.36.1	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.36.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-6.6.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.35.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-6.5.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.4.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.3.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.3.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.2.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.2.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.1.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.1.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.0.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-6.0.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-5.34.0	2.3.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.33.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.33.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.32.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.32.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.31.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.31.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.30.2	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.30.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.30.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
mr-5.29.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.28.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn
emr-5.28.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.27.1	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.27.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.26.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.25.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.24.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.24.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.23.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.23.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.22.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.21.2	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.21.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.21.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.20.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.20.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.19.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.19.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.18.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.18.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.17.2	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.17.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.17.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.16.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.16.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.15.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.15.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.14.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.14.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.14.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.13.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.13.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.12.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.12.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.12.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.12.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.11.4	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.11.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.11.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.11.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.11.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.10.1	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.10.0	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.9.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.9.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.8.3	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.8.2	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.8.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.8.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.7.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.7.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.6.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.6.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.4	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.5.3	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.5.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.4.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.4.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.3.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.3.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.3.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.2.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.2.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.2.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.2.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.1.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.1.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.0.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.0.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.0.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-5.0.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-4.9.6	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-4.9.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.9.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-4.9.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.9.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-4.9.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-4.8.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-4.8.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-4.8.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-4.7.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-4.7.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-4.6.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.6.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-4.5.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.4.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-4.3.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.2.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Étiquette de version Amazon EMR	Version de Hive	Composants installés avec Hive
emr-4.1.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.0.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Notes de mise à jour de Hive par version

Rubriques

- Amazon EMR 7.8.0 - Notes de mise à jour de Hive
- Amazon EMR 7.7.0 - Notes de mise à jour de Hive
- Amazon EMR 7.6.0 - Notes de mise à jour de Hive
- Amazon EMR 7.5.0 - Notes de mise à jour de Hive
- Amazon EMR 7.4.0 - Notes de mise à jour de Hive
- Amazon EMR 7.3.0 - Notes de mise à jour de Hive

- Amazon EMR 7.2.0 - Notes de mise à jour de Hive
- Amazon EMR 7.1.0 - Notes de mise à jour de Hive
- Amazon EMR 7.0.0 - Notes de mise à jour de Hive
- Amazon EMR 6.15.0 – Notes de mise à jour de Hive
- Amazon EMR 6.14.0 - Notes de mise à jour de Hive
- Amazon EMR 6.13.0 - Notes de mise à jour de Hive
- Amazon EMR 6.12.0 - Notes de mise à jour de Hive
- Amazon EMR 6.11.0 - Notes de mise à jour de Hive
- Amazon EMR 6.10.0 - Notes de mise à jour de Hive
- Amazon EMR 6.9.0 - Notes de mise à jour de Hive
- Amazon EMR 6.8.0 - Notes de mise à jour de Hive
- Amazon EMR 6.7.0 - Notes de mise à jour de Hive
- Amazon EMR 6.6.0 - Notes de mise à jour de Hive

Amazon EMR 7.8.0 - Notes de mise à jour de Hive

Amazon EMR 7.8.0 - Changements concernant Hive

Type	Description
Correctif de bogue.	Corrige CVE-2024-23953 : Apache Hive : attaque chronométrée contre une signature dans LLAP.

Problèmes connus

- Pour les requêtes Hive Insert Overwrite avec Amazon S3 Express One Zone comme emplacement de sortie, définissez la configuration du site principal : sur. `fs.s3a.directory.operations.purge.uploads` `false`

Amazon EMR 7.7.0 - Notes de mise à jour de Hive

Amazon EMR 7.7.0 - Changements concernant Hive

Type	Description
Correctif de bogue.	Corrige CVE-2024-29869 : Apache Hive : fichier d'informations d'identification créé avec des autorisations non restrictives.
Correctif de bogue.	SemanticException Corrige l'activation d'une politique de filtrage au niveau des lignes dans Apache Ran.
Correctif de bogue.	Désactivez Tez Async Init RR lorsque LLAP ou ACID est activé.

Problèmes connus

- Pour les requêtes Hive Insert Overwrite avec Amazon S3 Express One Zone comme emplacement de sortie, définissez la configuration du site principal : sur. `fs.s3a.directory.operations.purge.uploads` `false`

Amazon EMR 7.6.0 - Notes de mise à jour de Hive

Amazon EMR 7.6.0 - Changements concernant Hive

Type	Description
Amélioration	Ajout d'une fonction de listage rapide des préfixes S3 pour les tables partitionnées ORC non ACID
Fonctionnalité	Ajout du support pour les requêtes Magic Committers for Hive Write sur le système S3 AFile

Problèmes connus

- Pour les requêtes Hive Insert Overwrite avec Amazon S3 Express One Zone comme emplacement de sortie, définissez la configuration du site principal : sur. `fs.s3a.directory.operations.purge.uploads` `false`

Amazon EMR 7.6.0 - Nouvelles configurations

Classification	Nom	Par défaut	Description
hive-site	`hive.exec.fast.s3.partition.discovery.enabled`	true	S'il faut utiliser la découverte rapide des partitions S3 pour le calcul du fractionnement. Cela permettra d'obtenir une liste basée sur des préfixes pour les formats de fichiers pris en charge : ORC. Notez que cette fonctionnalité utilise un paramètre d'API S3 que la classe de stockage S3 Express One Zone ne prend pas en charge. Lorsque vous les utilisez, désactivez cette fonctionnalité.
hive-site	`hive.exec.fast.s3.partition.discovery.max.thread.threshold`	128	Le degré de parallélisme maximal pour une découverte rapide des partitions S3.

Classification	Nom	Par défaut	Description
hive-site	`hive.exec.fast.s3.partition.discovery.parallelism`	10	Degré de parallélisme d'une seule opération de découverte rapide de partitions S3. Cette configuration n'a d'effet que si elle `hive.exec.fast.s3.partition.discovery.enabled` est définie sur `true`

Classification	Nom	Par défaut	Description
hive-site	`hive.blobstore.output-committer.magic.track.commits.in.memory.enabled`	true	Drapeau pour activer Magic Committer with Hive devrait-il suivre tous les validations en attente en mémoire ? Le validateur Magic dispose d'une option permettant de stocker les données de validation en mémoire, ce qui peut accélérer l' TaskCommit opération en effectuant moins d'appels S3. Cette configuration remplace la configuration Hadoop `fs.s3a.committer.magic.track.commits.in.memory.enabled`

Classification	Nom	Par défaut	Description
hive-site	`hive.blobstore.output-committer.dp.skip.task.staging.dir.creation`	true	Drapeau pour activer Magic Committer doit-il créer les chemins de mise en scène DP dans le blobstore ? Cet indicateur n'est applicable que lors du suivi des validations en mémoire lorsque Hive utilise Magic Committer via. `hive.blobstore.output-committer.magic.track.commits.in.memory.enabled` Par défaut, il est défini sur true mais il ne prend effet que s'il `hive.blobstore.output-committer.magic.track.commits.in.memory.enabled` est activé et enregistre des appels S3 supplémentaires pour créer des chemins de tentative de tâche dans le blobstore.

Classification	Nom	Par défaut	Description
hive-site	`hive.blobstore.output-committer.magic.disable.fs.cache.for.llap`	true	Indicateur permettant d'activer si les caches Blobstore FS doivent être désactivés dans les flux d'écriture pour LLAP lors de l'utilisation de Magic Committer. Ce drapeau apparaît lorsque LLAP est activé et est défini par défaut sur true.

Amazon EMR 7.5.0 - Notes de mise à jour de Hive

Amazon EMR 7.5.0 - Changements concernant Hive

Type	Description
Amélioration	Le temps maximum d'attente pour l'ouverture de la session Tez pendant que vous essayez d'utiliser la session existante dans HiveCLI a été augmenté à 10 secondes
Amélioration	Configurations optimisées pour améliorer les performances dans les requêtes de sélection simples avec LIMIT

Amazon EMR 7.4.0 - Notes de mise à jour de Hive

Amazon EMR 7.4.0 - Changements concernant Hive

Type	Description
Upgrade	HIVE-28191 : mise à niveau de la version Hadoop vers la version 3.4.0
Upgrade	Mettre à jour le protobuf partagé de Hadoop vers la version 3.21
Upgrade	Mettre à jour commons-cli vers la version 1.5.0
Upgrade	Mettre à jour commons-compress vers la version 1.24.0
Upgrade	Mettre à jour commons-io vers la version 2.14.0
Upgrade	Mettre à jour commons-lang3 vers la version 3.21.0
Amélioration	Changez le temps d'attente pour l'ouverture de la session Tez tout en essayant d'utiliser la session existante dans HiveCLI à 10 secondes
Amélioration	Activez le mécanisme de court-circuit dans Tez DAG pour des requêtes de sélection simples avec LIMIT
Amélioration	HIVE-21100 : Autoriser l'aplatissement des sous-répertoires de table résultant de l'utilisa tion du moteur TEZ et de la clause UNION
Correctif de bogue.	HIVE-25095 : La commande Beeline/hive -e ne peut pas traiter une requête avec un guillemet
Correctif de bogue.	HIVE-13781 : Le Tez Job a échoué FileNotFo undException lorsque le répertoire de partition n'existe pas

Type	Description
Correctif de bogue.	HIVE-28480 : Désactiver l'incompatibilité du générateur de hachage SMB sur une partition entre les branches de jointure dans les versions précédentes de RS

Amazon EMR 7.4.0 - Nouvelles configurations

Classification	Nom	Par défaut	Description
hive-site	hive.ignore.failure.partition.dir .not.found	false	Ignore l'échec si la partition de table existe mais que le chemin de stockage réel de l'objet n'existe pas.
hive-site	hive.tez.union.flatten.subdirectories	false	Lorsque vous écrivez des données dans une table et que UNION ALL est la dernière étape de la requête, Hive on Tez crée un sous-répertoire pour chaque branche de UNION ALL. Lorsque cette propriété est activée, les sous-répertoires sont supprimés, les fichiers sont renommés et déplacés vers le répertoire parent. Notez que cela n'a

Classification	Nom	Par défaut	Description
			aucun effet lorsque hive.blobstore.use .output-committer est activé.

Amazon EMR 7.3.0 - Notes de mise à jour de Hive

Amazon EMR 7.3.0 - Changements concernant Hive

Type	Description
Fonctionnalité	HIVE-18728 — Web HCat sécurisé avec SSL.
Amélioration	Support de configuration des informations d'identification du keystore SSL pour l'interface utilisateur Web du démon LLAP.
Amélioration	Fournissez une option permettant de contrôler la vérification du nom d'hôte SSL pour le serveur de métastore Hive.
Correctif de bogue.	HIVE-26541 — NPE lors du démarrage du service Web. HCat
Correctif de bogue.	HIVE-23011 — L'optimiseur de travail partagé doit vérifier les prédicats résiduels lors de la comparaison des jointures.
Correctif de bogue.	Corrigez javax.security.sasl. SaslException: Aucune couche de protection commune entre le client et le serveur entre HMS et Namenode lorsque le chiffrement en transit est activé.
Correctif de bogue.	Corrigez l'IOExceptionendroit où la fin de la division des orcs se chevauche avec le début de l'emplacement d'un bloc.

Type	Description
Correctif de bogue.	Utilisez un délimiteur de nom de colonne au lieu de toujours le diviser par virgule lorsque les noms de colonne contiennent une virgule et d'utiliser. CSVSerde

Amazon EMR 7.3.0 - Nouvelles configurations

Classification	Nom	Par défaut	Description
hcatalog-webhcat-site	templeton.use.ssl	false	Définissez ce paramètre sur true pour utiliser le cryptage SSL pour le HCat serveur Web.
hcatalog-webhcat-site	templeton.keystore.path		Emplacement du keystore des certificats SSL pour le HCat serveur Web.
hcatalog-webhcat-site	templeton.keystore.password		Mot de passe du keystore du certificat SSL pour le HCat serveur Web.
hcatalog-webhcat-site	templeton.ssl.protocol.blacklist	SSLv2, SSLv3	Versions SSL à désactiver pour le HCat serveur Web.
hcatalog-webhcat-site	templeton.host	0.0.0.0	Adresse de l'hôte que le HCat serveur Web écoutera.

Classification	Nom	Par défaut	Description
hive-site	hive.metastore.ssl.enable.hostname.verification	false	Contrôlez la vérification du nom d'hôte lors de la prise de contact SSL/TLS.
hive-site	hive.llap.daemon.web.ssl.keystore.path		Emplacement du keystore du certificat SSL pour l'interface utilisateur Web du démon LLAP.
hive-site	hive.llap.daemon.web.ssl.keystore.password		Mot de passe du keystore du certificat SSL pour l'interface utilisateur Web du démon LLAP.
hive-site	hive.metastore.hadoop.rpc.protection.override .to.authentication	false	Lorsqu'il est activé, HMS remplace toujours la valeur de hadoop.rpc.protection pour l'authentification dans son ensemble de configurations.

Amazon EMR 7.2.0 - Notes de mise à jour de Hive

Amazon EMR 7.2.0 - Changements concernant Hive

Type	Description
Upgrade	Parquet 1.13.1 — Le parquet est mis à niveau vers la version 113.1.

Type	Description
Amélioration	HIVE-12930 — Support du shuffle SSL pour LLAP.
Amélioration	HIVE-23062 — Hive va vérifier l'URL de Yarn RM en mode TLS et Yarn HA pour une file d'attente Tez personnalisée.
Correctif de bogue.	HIVE-27952 — Hive ne parvient pas à créer SslContextFactory lorsqu'il possède plusieurs certificats. KeyStore
Correctif de bogue.	HIVE-28085 — Impossible d'accéder YarnQueueHelper à YARN compatible HTTPS. WebService
Correctif de bogue.	HIVE-26436 — Hive on MR NullPointerExcepti on lorsque InitializeOp n'a pas été appelé et que close a été appelé. Si l'opérateur n'a pas été initialisé, ignorez la fermeture.

Amazon EMR 7.2.0 - Nouvelles configurations

Classification	Nom	Par défaut	Description
hive-site	hive.llap.shuffle. ssl.enabled	false	Définissez ce paramètre sur true avec tez.runti me.shuffle.ssl.enable pour activer le shuffle SSL pour LLAP.

Amazon EMR 7.1.0 - Notes de mise à jour de Hive

Amazon EMR 7.1.0 - Changements concernant Hive

Type	Description
Correctif de bogue.	HIVE-24381 — La saisie de texte compressé renvoie 0 ligne si l'option ignorer l'en-tête/le pied de page est inclus.
Correctif de bogue.	HIVE-24190 — LLAP : ShuffleHandler peut renvoyer DISK_ERROR_EXCEPTION conformément à TEZ-4233.
Correctif de bogue.	HIVE-23073 — Shade Netty.
Correctif de bogue.	HIVE-23073 — Shade Netty et passe à Netty 4.1.48.Final.
Correctif de bogue.	HIVE-23148 — Le flux de clients externes Llap est interrompu à cause du netty shading.
Correctif de bogue.	HIVE-25180 — Améliore Netty.
Correctif de bogue.	HIVE-24524 — LLAP ShuffleHandler : passez à Netty4 et supprimez la dépendance Netty3 de Hive dans la mesure du possible.
Correctif de bogue.	HIVE-28000 — Hive QL : la clause « not in » donne des résultats incorrects lorsque la coercition de type ne peut pas avoir lieu.
Correctif de bogue.	HIVE-27993 — Netty4 ShuffleHandler doit utiliser 1 thread de boss.
Upgrade	Met à jour Netty vers la version 4.1.100.Final
Upgrade	Améliore Jetty vers la version 9.4.53.v2 0231009

Type	Description
Upgrade	Améliore Zookeeper vers la version 3.9.1

Amazon EMR 7.1.0 - Changements concernant Hive

- Amazon EMR 7.1 met à niveau Hive vers Netty 4.1.100.Final afin de résoudre les failles de sécurité de Netty3. Comme il dépend de netty3, Hive n' hive-druid-handlera pas le fichier `hive-druid-handler` JAR dans le chemin de classe de Hive dans Amazon EMR 7.1. Une prochaine version d'Amazon EMR l'inclura dans le classpath de Hive une fois que le gestionnaire Druid sera compatible avec les versions 4.1.100.Final ou ultérieures de Netty. Contactez l' AWS assistance si vous avez besoin du fichier `hive-druid-handler` JAR dans les versions 7.1 ou supérieures d'Amazon EMR.

Amazon EMR 7.0.0 - Notes de mise à jour de Hive

Amazon EMR 7.0.0 - Changements concernant Hive

Type	Description
Upgrade	Hive Runtime utilise désormais Java 17 par défaut. Reportez-vous au guide de mise à jour d'EMR 7.0.0 pour plus de détails.
Rétroportage	HIVE-17709 : supprimer les références Sun.Misc.Cleaner
Correctif de bogue.	Désactiver Tez Async Init RR lorsque LLAP ou ACID est activé

Amazon EMR 6.15.0 – Notes de mise à jour de Hive

Amazon EMR 6.15.0 – Changements concernant Hive

Type	Description
Fonctionnalité	Prise en charge pour TEZ-4397 – Pour l'ouverture fractionnée asynchrone Tez, Hive prend désormais en charge les configurations Tez décrites dans la rubrique Ouverture fractionnée asynchrone Tez.
Correctif de bogue	HIVE-25400 – Déplace la mise à jour du décalage dans `BytesColumnVector` vers `setValPreallocated` .
Correctif de bogue	HIVE-25190 – Corrige de nombreuses petites allocations dans `BytesColumnVector` .
Correctif de bogue.	Empaqueter des modules netty avec le serveur llap pour éviter les NoClassDefFoundexceptions lors du démarrage de LLapDaemon sur des nœuds de travail.
Upgrade	Permet de mettre à niveau Apache Hadoop vers la version 3.3.6.
Upgrade	HIVE-26684 – Mise à niveau de `maven-shade-plugin` vers la version 3.4.1.
Amélioration	Pour accélérer le démarrage du cluster Amazon EMR, réduisez le délai d'inactivité du script de démarrage HCatalog de 15 secondes.

Amazon EMR 6.14.0 - Notes de mise à jour de Hive

Amazon EMR 6.14.0 - Changements concernant Hive

Type	Description
Amélioration	HIVE-26762 : suppression de l'élagage des opérandes dans `HiveFilterSetOpTranposeRule`
Correctif de bogue	HIVE-27582 : Ne pas mettre en cache HBase le format d'entrée de la table dans FetchOperator
Correctif de bogue	HIVE-26452 : NPE lors de la conversion de JOIN en MAPJOIN et colonne JOIN référencée plusieurs fois
Correctif de bogue	HIVE-26416 : `AcidUtils.isRawFormatFile()` lance `InvalidProtocolBufferException` pour le fichier non-ORC
Correctif de bogue	HIVE-26105 : Afficher les colonnes affiche des valeurs supplémentaires si la colonne Commentaires contient un caractère chinois spécifique
Correctif de bogue	HIVE-25864 : l'optimisation des requêtes Hive crée un plan erroné pour le transfert des prédicats avec fonction de fenêtrage
Correctif de bogue	HIVE-25224 : plusieurs instructions INSERT impliquant des tables avec différentes `bucketing_versions` renvoient des messages d'erreur
Correctif de bogue	HIVE-24151 : `MultiDelimitSerDe` déplace les données si les chaînes contiennent des caractères non-ASCII

Type	Description
Correctif de bogue	HIVE-23606 : (LLAP) retard dans le nettoyage de `DirectByteBuffer` pour `EncodedReaderImpl`
Correctif de bogue	HIVE-22165 : la synchronisation introduite par HIVE-14296 sur `SessionManager.closeSession` entraîne une latence élevée sur un serveur Hive très fréquenté
Correctif de bogue	HIVE-21304 : renforcement de l'utilisation de la version de compartimentation

Amazon EMR 6.13.0 - Notes de mise à jour de Hive

Amazon EMR 6.13.0 - Changements concernant Hive

Type	Description
Amélioration	Mettre à niveau les scripts Python pour qu'ils prennent en charge Python3
Amélioration	HIVE-27097 : Améliorer la stratégie de nouvelle tentative pour le client et le serveur MetaStore
Correctif de bogue.	HIVE-21778 : CBO : « La structure n'est pas nulle » est considérée comme nulle, ce qui entraîne toujours l'absence de filtre dans la requête
Correctif de bogue.	HIVE-21009 : ajout de la possibilité pour l'utilisateur de définir un utilisateur de liaison
Correctif de bogue.	HIVE-22661 : le compactage échoue sur une table non compartimentée avec des données chargées dans le chemin

Type	Description
Correctif de bogue.	HIVE-19718 : l'ajout de partitions en masse permet également d'extraire une table pour chaque partition
Correctif de bogue.	HIVE-22173 : la requête avec plusieurs vues latérales se bloque pendant la compilation
Correctif de bogue.	HIVE-27088 : résultats incorrects lorsque les jointures internes et externes avec des filtres de post-jointure sont fusionnées
Correctif de bogue.	HIVE-21935 : vectorisation de Hive : performan ces dégradées avec la vectorisation UDF
Correctif de bogue.	HIVE-25299 : la conversion de l'horodatage en types de données numériques est incorrecte pour les fuseaux horaires autres que l'UTC
Correctif de bogue.	HIVE-24626 : LLAP : les fils de lecture peuvent être privés si tous les threads de l'ascense ur d'E/S sont occupés pour être mis en file d'attente vers un autre lecteur avec une file d'attente complète
Correctif de bogue.	HIVE-27029 : la requête hive échoue avec une erreur de fermeture du système de fichiers, refonte effectuée pour HIVE-26352
Correctif de bogue.	HIVE-26352 : la vérification de l'accès à la file d'attente Tez échoue avec une exception GSS lors du compactage
Correctif de bogue.	HIVE-24590 : la journalisation des opérations divulgue toujours les appendices log4j

Type	Description
Correctif de bogue.	HIVE-24552 : Possibilité de fuite ou d'accumulation de connexions HMS dans loadDynamicPartitions
Correctif de bogue.	HIVE-27069 : résultats incorrects avec la jointure par mappage de compartiment
Correctif de bogue.	HIVE-27344 : Ajouter une vérification nulle dans #close RecordReaderImpl
Correctif de bogue.	HIVE-27439 : prise en charge de l'espace en décimal
Correctif de bogue.	HIVE-27267 : résultats incorrects lors de la jointure par mappage de compartiment sur une colonne décimale avec sous-requête
Correctif de bogue.	HIVE-21986 : Interface utilisateur HiveServer Web : définition de l'en-tête de réponse par Strict-Transport-Security défaut
Correctif de bogue.	HIVE-22148 : les jetons de délégation S3A ne sont pas ajoutés dans la configuration des tâches du compacteur.
Correctif de bogue.	HIVE-22622 : Hive permet de créer une structure avec des noms d'attributs dupliqués
Correctif de bogue.	HIVE-22008 : l'opérateur LIKE doit correspondre à une entrée multiligne
Correctif de bogue.	HIVE-23144 : LLAP : Laisse le nettoyage sur ServiceStop QueryTracker
Correctif de bogue.	HIVE-22391 : NPE lors de la vérification du cache des résultats des requêtes Hive

Type	Description
Correctif de bogue.	HIVE-23305 : NullPointerException dans LlapTaskSchedulerService AddNode en raison des conditions de course
Correctif de bogue.	HIVE-22178 : Parquet jeté après FilterPredicate CastException SchemaEvolution
Correctif de bogue.	HIVE-21517 : Corriger AggregateStatsCache
Correctif de bogue.	HIVE-21825 : Améliore le message d'erreur du client lorsque la haute disponibilité active/passive est activée
Correctif de bogue.	HIVE-23389 : FilterMergeRule peut conduire à AssertionError
Correctif de bogue.	HIVE-22767 : Beeline n'analyse pas correctement les points-virgules dans les commentaires
Correctif de bogue.	HIVE-22996 : l' BasicStats analyse doit vérifier de manière proactive la présence d'une chaîne nulle ou vide
Correctif de bogue.	HIVE-22808 : ne gère pas HiveRelFieldTrimmer HiveTableFunctionScan
Correctif de bogue.	HIVE-22437 : le cache de métadonnées LLAP NPE lors du verrouillage des métadonnées.
Correctif de bogue.	HIVE-22606 : AvroSerde enregistre avro.schema.literal au niveau INFO
Correctif de bogue.	HIVE-22713 : une propagation constante ne doit pas être effectuée pour la structure Join-Fil(*)-RS

Type	Description
Correctif de bogue.	HIVE-21624 : LLAP : les métriques du processeur au niveau du thread sont cassées
Correctif de bogue.	HIVE-22815 : réduisez la création inutile d'objets de système de fichiers dans MROutput
Correctif de bogue.	HIVE-23060 : échec de la requête avec erreur « Grouping sets expression is not in GROUP BY key. » Une erreur s'est produite à proximité du jeton »
Correctif de bogue.	HIVE-22236 : impossible de créer une vue en sélectionnant la vue contenant une sous-requ ête `NOT IN`
Correctif de bogue.	HIVE-19886 : Les journaux peuvent être dirigés vers 2 fichiers si -hiveconf hive.log.file est utilisé.
Correctif de bogue.	HIVE-20620 : collisions de manifest lors de l'insertion dans des tables MM triées par compartiments avec partitionnement dynamique
Correctif de bogue.	HIVE-14557 : Nullpointer lorsque Mapjoin et Mapjoin sont activés SkewJoin
Correctif de bogue.	HIVE-20471 : problèmes pour obtenir le chemin de base de données par défaut
Correctif de bogue.	HIVE-20598 : Corriger les fautes de frappe dans les calculs HiveAlgorithmsUtil
Correctif de bogue.	HIVE-14737 : problème lors de l'accès à /logs dans une interface utilisateur Web de Hive Server 2 activée pour Kerberos

Type	Description
Correctif de bogue.	HIVE-20733 : Le UDFOPEqual NS générique ne peut pas utiliser = dans les descriptions des plans
Correctif de bogue.	HIVE-20848 : Après la configuration, la UpdateInputAccessTimeHook requête échoue avec Table Not Found.
Correctif de bogue.	HIVE-18929 : La méthode humanReadableInt du HiveStringUtils fichier .java comporte une condition de race.
Correctif de bogue.	HIVE-20841 : LLAP : rendre les ports dynamiques configurables
Correctif de bogue.	HIVE-20930 : VectorCoalesce en mode FILTRE, cela ne prend pas effet
Correctif de bogue.	HIVE-21007 : Semi-join + Union peut mener à de mauvais plans
Correctif de bogue.	HIVE-21074 : l'élimination des requêtes dans les tables compartimentées de Hive ne fonctionne pas pour la condition IS NOT NULL.
Correctif de bogue.	HIVE-21223 : CachedStore renvoie une partition nulle lorsque la partition n'existe pas
Correctif de bogue.	HIVE-19625 : NPE potentiel et masquage d'une exception réelle dans Hive#copyFiles
Correctif de bogue.	HIVE-17020 : Un dédup RS agressif peut supprimer de manière incorrecte une branche d'arbre OP
Correctif de bogue.	HIVE-20168 : Journalisation masquée ReduceSinkOperator

Type	Description
Correctif de bogue.	HIVE-20879 : L'utilisation de null dans une expression de projection entraîne CastExcep tion
Correctif de bogue.	HIVE-20888 TxnHandler : sort () appelé sur des listes immuables
Correctif de bogue.	HIVE-19948 : HiveCli ne divise pas correctem ent la commande par point-virgule si des guillemets sont à l'intérieur de la chaîne
Correctif de bogue.	HIVE-20621 : GetOperationStatus appelé dans resultset.next provoquant une lenteur incrément ielle
Correctif de bogue.	HIVE-20854 : Valeurs par défaut raisonnables : l'intervalle de battement de cœur Zookeeper de Hive est de 20 minutes, passer à 2
Correctif de bogue.	HIVE-20330 : Le HCat chargeur ne peut pas gérer plusieurs InputJobInfo objets pour une tâche comportant plusieurs entrées
Correctif de bogue.	HIVE-20787 : Le boîtier MapJoinBytesTableC ontainer DummyRow ne gère pas la réutilisa tion
Correctif de bogue.	HIVE-20331 : la requête avec union all, lateral view et Join échoue avec « cannot find parent in the child operator »
Correctif de bogue.	HIVE-19968 : l'exception UDF n'est pas déclenchée
Correctif de bogue.	HIVE-20410 : l'annulation de l'insertion et de la réécriture sur la table transactionnelle provoque l'erreur « Not enough history available for… »

Type	Description
Correctif de bogue.	HIVE-20059 : le streaming Hive devrait essayer le préfixe shade sans condition, sauf exception
Correctif de bogue.	HIVE-19424 : Entrée NPE MetaDataFormatters
Correctif de bogue.	HIVE-20355 : Paramètre de nettoyage de .setSchema HiveConnection
Correctif de bogue.	HIVE-20858 : Le sérialiseur n'est pas correctement initialisé avec la configuration dans Utilities. createEmptyBuckets
Correctif de bogue.	HIVE-20424 : schematool ne doit pas polluer l'historique de Beeline
Correctif de bogue.	HIVE-20338 : LLAP : identifiant de fichier synthétique forcé pour les systèmes de fichiers dotés du protocole HDFS avec une sémantique de mutation POSIX
Correctif de bogue.	HIVE-11708 : Les opérateurs logiques se déclenchent avec NULL ClassCastExceptions
Correctif de bogue.	HIVE-21082 : dans HPL/SQL, l'instruction declare ne prend pas en charge les variables de type caractère
Correctif de bogue.	HIVE-16690 : configuration de la limite du produit cartésien Tez en fonction de la taille du cluster LLAP
Correctif de bogue.	HIVE-14516 :. OrcInputFormat SplitGenerator. Appel interne
Correctif de bogue.	HIVE-20981 : streaming/fuites AbstractR ecordWriter HeapMemoryMonitor

Type	Description
Correctif de bogue.	HIVE-20043 : HiveServer 2 : SessionState possède un bloc de synchronisation statique autour d'un AtomicBoolean
Correctif de bogue.	HIVE-20191 : l'application du PreCommit correctif n'échoue pas si le correctif est vide
Correctif de bogue.	HIVE-20400 : la création d'une table doit toujours utiliser un chemin entièrement qualifié pour éviter toute ambiguïté potentielle avec FS
Correctif de bogue.	Ajoutez une vérification nulle pour skewedInfo avant d'accéder aux colonnes asymétriques

Amazon EMR 6.12.0 - Notes de mise à jour de Hive

Amazon EMR 6.12.0 - Changements concernant Hive

Type	Description
Amélioration	Ajout de la prise en charge de l'exécution JDK 11 et JDK 17
Amélioration	Ajoutez la possibilité de demander des noms de colonnes de mots clés réservés et sensibles aux majuscules lors de l'utilisation de S3 Select. Pour l'utiliser, définissez la propriété de la table au format « s3select.column.mapping » = "*column1*:*fieldName1* , *column2*:*fieldName2* ,..."
Amélioration	HIVE-23133 : les opérations numériques peuvent avoir des résultats différents selon les architectures matérielles

Type	Description
Amélioration	HIVE-27145 : À utiliser StrictMath pour les fonctions mathématiques restantes en tant que suivi de HIVE-23133
Correctif de bogue.	Correction de l'incompatibilité des caractères génériques dans get_partitions_by_filter et get_num_partitions_by_filter HMS causée par le portage du HIVE-22900 dans EMR Hive 6.4.0 APIs
Correctif de bogue.	HIVE-26736 : échec de l'autorisation pour les vues imbriquées comportant une clause WITH
Correctif de bogue.	HIVE-22416 : les journaux des opérations liés à MR sont manquants lorsque l'exécution parallèle est activée
Correctif de bogue.	HIVE-19653 : pushdown de prédicat incorrect pour groupby avec des ensembles de regroupement
Correctif de bogue.	HIVE-22094 : échec des requêtes avec ClassCastException : hive.ql.exec.vector. DecimalColumnVectorne peut pas être converti en Hive.QL.Exec.Vector.Decimal64 ColumnVector
Correctif de bogue.	HIVE-26340 : l'opérateur PTF vectorisé échoue si la requête comporte une fonction window en majuscules
Correctif de bogue.	HIVE-26184 : COLLECT_SET avec GROUP BY est très lent lorsque certaines clés sont très asymétriques

Type	Description
Correctif de bogue.	HIVE-26373 : ClassCastException lors de la lecture des horodatages d'une table contenant des données Avro HBase
Correctif de bogue.	HIVE-26388 : ClassCastException lorsqu'il y a une colonne autre que de type chaîne dans la table source de la requête CTAS Upgrade HIVE-26172 : Hive - Upgrade Ant vers la version 1.10.11 en raison de CVE-2021-36373 et CVE-2021-36374
Correctif de bogue.	HIVE-26114 : corriger la connexion jdbc (hiveserver2) en utilisant la commande dfs avec un espace de préfixe provoquera une exception
Correctif de bogue.	HIVE-26396 : la fonction trunc a un problème d'interception de précision et le résultat a de nombreux 0
Correctif de bogue.	HIVE-26446 : Impossible de renseigner le champ HiveProtoLoggingHook pour les tables partitionnées. TablesWritten
Correctif de bogue.	HIVE-26639 : ConstantVectorExpression et ExplainTask ne devrait pas s'appuyer sur le jeu de caractères par défaut
Correctif de bogue.	HIVE-22670 : ArrayIndexOutOfBoundsException lorsqu'un lecteur vectorisé est utilisé pour lire un dossier de parquet
Correctif de bogue.	HIVE-23607 : Problème d'autorisation : la création d'une vue sur une autre vue réussit mais la modification de la vue échoue

Type	Description
Correctif de bogue.	HIVE-25498 : une requête comportant plus de 31 fonctions « count distinct » renvoie un résultat erroné
Correctif de bogue.	HIVE-25780 : DistinctExpansion crée plus de 64 ensembles de regroupement II
Correctif de bogue.	HIVE-23868 : Spécification de la fonction de fenêtrage : support 0 précédant/suivant
Correctif de bogue.	HIVE-24539 : la génération du OrcInputFormat schéma doit respecter le délimiteur de colonne
Correctif de bogue.	HIVE-23476 : LAP : préalloue également des arènes pour le cas de mmap
Correctif de bogue.	HIVE-25806 : Fuite possible LlapCache AwareFs - Parquet, LLAP IO
Correctif de bogue.	HIVE-23498 : Désactiver la méthode HTTP Trace sur ThriftHttpCliService
Correctif de bogue.	HIVE-25729 : ThriftUnionObjectInspector doit être notifié lorsqu'il est complètement activé
Correctif de bogue.	HIVE-23846 : Évitez la sérialisation et la désérialisation inutiles des vecteurs de bits
Correctif de bogue.	HIVE-24233 : sauf que la sous-requête lance un pointeur nul avec cbo désactivé
Correctif de bogue.	HIVE-24276 : Vulnérabilité liée à HiveServer 2 scripts intersites (XSS) dans loggerconf jsp
Correctif de bogue.	HIVE-25721 : Le résultat de la jointure externe est erroné

Type	Description
Correctif de bogue.	HIVE-25223 : Sélectionner avec limite ne renvoie aucune ligne sur une table non native
Correctif de bogue.	HIVE-25794 : CombineHiveRecordReader : les instructions de log dans une boucle entraînent une pression sur la mémoire
Correctif de bogue.	HIVE-23602 : Utiliser le package Java simultané pour le set de descripteurs d'opérati ons
Correctif de bogue.	HIVE-24045 : Aucune journalisation liée à la création de la base de données par défaut
Correctif de bogue.	HIVE-24305 : le schéma décimal avro ne remplit pas correctement l'échelle/la précision si la valeur est entre guillemets
Correctif de bogue.	HIVE-25040 : Supprimer la cascade de bases de données ne peut pas supprimer les fonctions persistantes
Correctif de bogue.	HIVE-23501 : AOOB VectorDeserializeRow lorsque les types complexes sont convertis en types primitifs
Correctif de bogue.	HIVE-23704 : Le serveur HTTP Thrift ne gère pas correctement le gestionnaire d'authent ification
Correctif de bogue.	HIVE-23529 : Le CTAS est interrompu pour uniontype lorsque row_deserialize
Correctif de bogue.	HIVE-24144 : La getIdentifierQuote chaîne renvoie une valeur incorrecte HiveDatab aseMetaData

Type	Description
Correctif de bogue.	HIVE-23850 : Autoriser PPD lorsque le sujet n'est pas une colonne avec des ensembles de regroupement présents
Correctif de bogue.	HIVE-25919 : ClassCastException lorsque vous insérez un prédicat de colonne booléen HBase StorageHandler
Correctif de bogue.	HIVE-25261 : La nouvelle tentative HMSHandle r devrait inclure une brève description MetaException de la cible
Correctif de bogue.	HIVE-24792 : Fuite de fil potentielle en cours d'exploitation
Correctif de bogue.	HIVE-23409 : Si la réouverture de TezSessio n l'application échoue en raison de la panne du service Timeline, le formulaire par défaut SessionPool est fermé TezSession après une nouvelle tentative
Correctif de bogue.	HIVE-23615 : Ne pas déférencer les pointeurs nuls dans la classe de commandes Beeline
Correctif de bogue.	HIVE-24849 : Création d'un délai d'expiration du socket de table externe lorsque l'emplace ment contient un grand nombre de fichiers (affecte la version 3.1.2)
Correctif de bogue.	HIVE-25209 : Requête SELECT avec fonction SUM produisant un résultat inattendu
Correctif de bogue.	HIVE-23666 : checkHashMode L'efficacité est ignorée lorsqu'un opérateur groupby n'a pas d'ensemble de regroupement

Type	Description
Correctif de bogue.	HIVE-23873 : L'interrogation de la table Hive JDBCStorage Handler échoue avec NPE lorsque le CBO est désactivé
Correctif de bogue.	HIVE-24149 : HiveStreamingConnection ne ferme pas la connexion HMS
Correctif de bogue.	HIVE-25561 : La tâche supprimée ne doit pas valider le fichier. (affecte les versions 2.x et 3.x)
Correctif de bogue.	HIVE-25683 : Fermez le lecteur. AcidUtils isRawFormatDossier
Correctif de bogue.	HIVE-24294 : TezSessionPool les sessions peuvent être lancées AssertionError
Correctif de bogue.	HIVE-24182 : Problème d'autorisation de ranger avec un permanent UDFs
Correctif de bogue.	HIVE-22805 : La vectorisation avec un tableau ou une map conditionnelle n'est pas implément ée et provoque une erreur
Correctif de bogue.	HIVE-22828 : Decimal64 : les instructions NVL et CASE convertissent implicitement decimal64 en 128
Correctif de bogue.	HIVE-21398 : Les colonnes contenant des statistiques estimées ne doivent pas être considérées comme des clés uniques
Correctif de bogue.	HIVE-22490 : L'ajout de fichiers jar contenant des caractères spéciaux dans leur chemin génère une erreur

Type	Description
Correctif de bogue.	HIVE-22700 : Les compactages peuvent entraîner une fuite de mémoire lorsqu'ils ne sont pas autorisés
Correctif de bogue.	HIVE-22053 : Le nom de la fonction n'est pas normalisé lors de la création de la fonction
Correctif de bogue.	HIVE-22595 : Les insertions de partitions dynamiques échouent sur la table Avro avec schéma externe
Correctif de bogue.	HIVE-21795 : La ligne récapitulative du cumul peut être manquante lorsqu'un mapjoin se produit sur une table partitionnée
Correctif de bogue.	HIVE-22987 : ClassCastException dans when est nul VectorCoalesce DataTypePhysicalVariation
Correctif de bogue.	HIVE-22219 : La désactivation d'un gestionnaire de nœuds bloque le redémarrage du service LLAP
Correctif de bogue.	HIVE-21793 : CBO récupère les statistiques des colonnes même si hive.stats.fetch.column.stats est défini sur false
Correctif de bogue.	HIVE-22163 : CBO : l'activation du CBO active l'estimation des statistiques, même lorsque l'estimation est désactivée
Correctif de bogue.	HIVE-18735 : Create table like perd l'attribut transactionnel

Type	Description
Correctif de bogue.	HIVE-22433 : Gestionnaire de stockage Hive JDBC : résultats incorrects extraits de BOOLEAN et TIMESTAMP de la source de données JDBC DataType
Correctif de bogue.	HIVE-19430 :. ObjectStore cleanNotificationEvents OutOfMemory sur un grand nombre d'événements en attente
Correctif de bogue.	HIVE-20785 : Nom de clé incorrect dans le JDBC. DatabaseMetaData getPrimaryKeys méthode
Correctif de bogue.	HIVE-16116 : Beeline lance un NPE lorsque beeline.hiveconfvariables={} dans beeline.properties
Correctif de bogue.	HIVE-20066 : hive.load.data.owner est comparé au principal complet
Correctif de bogue.	HIVE-20489 : Explain plan of query se bloque
Correctif de bogue.	HIVE-21033 : L'oubli de fermer l'opération coupe encore 2 sorties HiveServer
Correctif de bogue.	HIVE-19888 : Avertissement trompeur « METASTORE_FILTER_HOOK sera ignoré » de SessionState
Correctif de bogue.	HIVE-20303 : INSERT OVERWRITE TABLE db.table PARTITION (...) S'IL N'EXISTE PAS, lance InvalidTableException
Correctif de bogue.	HIVE-16144 : CompactionInfo n'a pas Equals/HashCode mais est utilisé dans Set

Type	Description
Correctif de bogue.	HIVE-20818 : Les vues créées avec une sous-requête WHERE considéreront les vues référencées dans la sous-requête comme une entrée directe
Correctif de bogue.	HIVE-21005 : LLAP : lecture d'un plus grand nombre de bandes par fuite ZlibCodecs
Correctif de bogue.	HIVE-20771 : LazyBinarySerDe échoue sur les structures vides.
Correctif de bogue.	HIVE-18852 : Message d'erreur trompeur lors de la validation de la table modifiée
Correctif de bogue.	HIVE-21124 : HPL/SQL ne supporte pas l'instruction CREATE TABLE LIKE
Correctif de bogue.	HIVE-20935 : Le téléchargement de l'archive tar du package LLAP échoue, ce qui entraîne un échec du démarrage du service LLAP EC2
Correctif de bogue.	HIVE-20409 : Hive ACID : Update/delete/merge ne nettoie pas le répertoire de mise en scène HDFS
Correctif de bogue.	HIVE-20570 : Union ALL avec hive.optimize.union.remove=true a un plan incorrect
Correctif de bogue.	HIVE-20421 : Entité de caractère « \b » illégale dans hive-default.xml.template
Correctif de bogue.	HIVE-19133 : Les indicateurs de performance par phase de HS2 WebUI ne s'affichent pas correctement
Correctif de bogue.	HIVE-18977 : La liste des partitions renvoie des résultats différents avec JDO et Direct SQL

Type	Description
Correctif de bogue.	HIVE-20034 : Annuler les modifications apportées à la gestion des MetaStore exception s à des fins de rétrocompatibilité
Correctif de bogue.	HIVE-20672 : Le thread de connexion LlapTaskSchedulerService doit signaler chaque intervalle fixe
Correctif de bogue.	HIVE-12812 : Activez mapred.input.dir.r ecursive par défaut pour prendre en charge l'union avec la fonction d'agrégation
Correctif de bogue.	HIVE-20147 : L'ingestion du streaming Hive est satisfaite de la journalisation synchronisée
Correctif de bogue.	HIVE-19203 : Problème de sécurité des threads dans HiveMetaStore
Correctif de bogue.	HIVE-20091 : Tez : ajouter des informations de sécurité pour la sortie FileSinkOperator
Correctif de bogue.	HIVE-16906 : Hive ATSHook doit vérifier la présence de yarn.timeline-service.enabled avant de se connecter à ATS
Correctif de bogue.	HIVE-20714 : SHOW tblproperties pour une seule propriété renvoie la valeur dans la colonne du nom
Correctif de bogue.	HIVE-24730 : Les classes Shims remplacent silencieusement les valeurs de hive-site.xml et tez-site.xml
Correctif de bogue.	HIVE-22055 : le nombre sélectionné donne un résultat incorrect après le chargement des données depuis un fichier texte

Amazon EMR 6.11.0 - Notes de mise à jour de Hive

Amazon EMR 6.11.0 - Changements concernant Hive

Type	Description
Amélioration	Ajout de la prise en charge de la suppression multithread de partitions afin d'améliorer les performances de suppression de partitions
Amélioration	Support de lecture des fichiers de requête Hive codés
Amélioration	Gestionnaire Tez Shuffle activé par défaut pour les tâches Hive on Tez
Bogue	Ajout d'une option pour activer la distribution déterministe des clés aux réducteurs afin de corriger un résultat incorrect lorsque hive.groupby.skewindata est activé (rapporté dans HIVE-20220)
Bogue	Correction d'un échec du calcul des statistiques lorsque le nom de partition par défaut est configuré
Bogue	Respectez tous les paramètres de classification SSL personnalisés transmis lorsque le protocole SSL est configuré prêt à l'emploi pour HiveServer 2 dans un cluster avec le chiffrement en transit activé
Rétroportage	HIVE-23617 : Problèmes d'API de stockage résolus FindBug
Rétroportage	HIVE-26408 : Vectorisation : corrigez la désallocation des colonnes à gratter, ne réutilisez pas un enfant comme sortie ConstantVectorExpression

Type	Description
Rétroportage	HIVE-23614 : Passez toujours à HiveConfig removeTempOr DuplicateFiles
Rétroportage	HIVE-23354 : Supprimer la vérification de la taille du fichier dans compareTempOr Duplicate Files
Rétroportage	HIVE-20344 : Corrigé PrivilegeSynchronizer pour le lancer de SBA. AccessControlExcep tion La propriété hive.privilege.synchronizer a également été introduite pour désactiver le synchronisateur de privilèges
Rétroportage	HIVE-15826 : Support de la configuration de « serialization.encoding » pour tous SerDes
Rétroportage	HIVE-18284 : Corrige le NPE lors de l'inserti on de données avec la clause « distribute by » avec optimisation du tri Dynpart
Rétroportage	HIVE-24930 : Le court-circuit Operator. setDone() provenant de l'opération enfant n'est pas utilisé dans le chemin de code vectorisé (si ChildSize == 1)
Rétroportage	HIVE-24523 : Le chemin de lecture vectorisé pour LazySimpleSerde ne respecte pas le SERDEPROPERTIES pour l'horodatage
Rétroportage	HIVE-23265 : Les ensembles de lignes dupliqués sont renvoyés avec les paramètres Limit et Offset définis
Rétroportage	HIVE-21492 : VectorizedParquetRecordReader impossible de lire le fichier de parquet généré à l'aide d'un outil économique/personnalisé

Type	Description
Rétroportage	HIVE-22540 : Vectorisation : les colonnes Decimal64 ne fonctionnent pas avec. Vectorize dBatchUtil makeLikeColumnVecteur ()
Rétroportage	HIVE-22588 : Vider les lignes restantes pour le reste des ensembles de regroupement lors du changement de mode groupby vectoriel
Rétroportage	HIVE-22551 : BytesColumnVector InitBuffe r doit nettoyer le vecteur et la longueur de manière cohérente
Rétroportage	HIVE-22448 : CBO : étendre le nombre multiple de manière distincte à l'aide d'une clé de regroupement
Rétroportage	HIVE-22248 : Corriger les problèmes persistan ts liés aux statistiques
Rétroportage	HIVE-22210 : La vectorisation peut réutiliser les colonnes de sortie de calcul impliquées dans le filtrage
Rétroportage	HIVE-21531 : Vectorisation : tous les hashcodes NULL ne sont pas calculés avec Murmur3
Rétroportage	HIVE-20419 : Vectorisation : empêche la mutation de VectorPartitionDesc après son utilisation dans une clé de hachage
Rétroportage	HIVE-19388 : lors de l'initialisation ClassCast Exception VectorMapJoinCommonOperator
Rétroportage	HIVE-21584 : Préparation à Java 11 : le chargeur de classe système n'est pas un chargeur URLClass

Type	Description
Rétroportage	HIVE-25107 : La journalisation des chemins de classe doit être au niveau DEBUG (#2271)
Rétroportage	HIVE-22097 : java.util incompatible. ArrayList pour Java 11
Rétroportage	HIVE-23938 : LLAP : JDK11 - certains arguments jvm liés à la rotation des fichiers journaux GC ne peuvent plus être utilisés
Rétroportage	HIVE-26226 : Exclure jdk.tools dep de hive-metastore dans upgrade-acid
Rétroportage	HIVE-17879 : mettre à jour le plug-in Datanucleus Maven
Rétroportage	HIVE-27004 : DateTimeFormatterBuilder # appendZoneText ne peut pas analyser 'UTC+' dans les versions Java supérieures à 8
Rétroportage	HIVE-16812 : VectorizedOrcAcidRowBatchReader ne filtre pas les événements de suppression
Rétroportage	RUCHE 17917 :. VectorizedOrcAcidRowBatchReader computeOffsetAndOptimisation du godet
Rétroportage	HIVE-19985 : ACID : ignore le décodage des sections ROW__ID pour les requêtes en lecture seule
Rétroportage	HIVE-20635 : VectorizedOrcAcidRowBatchReader ne filtre pas les événements de suppression pour les fichiers originaux
Upgrade	Mettre à niveau Javadoc vers 3.3.1

Type	Description
Upgrade	Mettre à niveau Javassist vers la version 3.24.1-GA
Upgrade	Mise à jour apache-directory-server en 2.0.0-M14

Nouvelles configurations

Nom	Classification	Description
hive.metastore.fs.drop.partition.threads	hive-site	Nombre de threads principaux dans le pool de threads de la partition de dépôt.
hive.metastore.fs.drop.partition.keepalive.time	hive-site	Durée en secondes pendant laquelle un thread asynchrone de partition inactif (issu du pool de threads) attendra l'arrivée d'une nouvelle tâche avant de se terminer.
hive.metastore.fs.drop.partition.threadpool.max.queue.size	hive-site	Taille de file d'attente maximale à utiliser dans le pool de threads pour supprimer des partitions du système de fichiers.
hive.groupby.enable.deterministic.distribution	hive-site	Activez la distribution déterministe des clés aux réducteurs. Il transmettra une valeur initiale constante lors de l'appel de la fonction rand utilisée pour le partitionnement aléatoire.

Nom	Classification	Description
hive.privilege.synchronizer	hive-site	S'il faut synchroniser périodiquement les privilèges de l'autorisateur externe en HiveServer 2.
hive.cli.query.file.encoding	hive-site	Codage de fichier pour tous les types de fichiers de requête (fichier de requête, fichier de requête d'initialisation, fichier rc, etc.) fournis dans les arguments cli.
hive.emr.tez.shuffle.enabled	hive-site	Les jobs Hive on Tez utilisent désormais tez_shuffle par défaut au lieu de mapreduce_shuffle comme gestionnaire de shuffle par défaut.

Configurations obsolètes

Les propriétés de configuration suivantes sont obsolètes en raison de HIVE-23354 et ne sont plus prises en charge avec les versions 6.11.0 et ultérieures d'Amazon EMR.

Nom	Valeur par défaut
`hive.mapred.reduce.tasks.speculative.execution`	`false`
`tez.am.speculation.enabled`	`false`

Amazon EMR 6.10.0 - Notes de mise à jour de Hive

Amazon EMR 6.10.0 - Changements concernant Hive

Type	Description
Fonctionnalité	Activez les contrôles d'accès AWS Lake Formation basés pour les requêtes Apache Hive (écriture) via IAM Passthrough (API Hivecli/Steps).
Amélioration	Désactiver la configuration `hive.log.explain.output` par défaut pour réduire la taille du journal
Rétroportage	HIVE-26408 : Vectorisation : corrigez la désallocation des colonnes à gratter, ne réutilisez pas un enfant comme sortie ConstantVectorExpression
Rétroportage	HIVE-22269 : Corrige le nombre de réducteur s erronés dans les requêtes d'insertion avec partition dynamique en raison de l'absence de statistiques causée par HIVE-20703.
Rétroportage	HIVE-22891 : Ignorer l' PartitionDesc extraction CombineHiveRecord pour le mode d'exécution non LLAP
Rétroportage	HIVE-23804 : Ajouter une base de données par défaut pour les tables spécifiques aux statistiq ues des colonnes dans le schéma de métastore Hive pour les rendre rétrocompatibles
Rétroportage	HIVE-25277 : Suppression lente de la partition Hive pour les magasins d'objets Cloud, ce qui coûte cher ListFiles

Type	Description
Rétroportage	HIVE-19202 : Le CBO a échoué à cause d'un in. NullPointerException HiveAggregate isBucketedInput()
Rétroportage	HIVE-19048 : Fix Beeline : les erreurs Initscript sont ignorées
Rétroportage	HIVE-21085 : Le registre des vues matérialisées démarre une session tez non externe
Rétroportage	HIVE-21675 : CREATE VIEW IF NOT EXISTS renvoie une erreur plutôt que « OK » si la vue existe déjà. Il s'agit d'une régression par rapport à Hive 2.
Rétroportage	HIVE-21646 : Tez : Empêche d'échapper au contexte TezTasks de journalisation des threads
Rétroportage	HIVE-22054 : Évitez les listes récursives pour vérifier si un répertoire est vide
Rétroportage	HIVE-16587 : NPE lors de l'insertion de types complexes avec des valeurs nulles imbriquées
Rétroportage	HIVE-22647 : Activer le pool de sessions par défaut
Rétroportage	HIVE-13288 : Message d'exception confus dans .LocalizeResource DagUtils
Rétroportage	HIVE-23870 : Optimisez plusieurs conversions de texte dans. WritableHiveCharObjectInspector getPrimitiveJavaObjet/ HiveCharWritable
Rétroportage	HIVE-21498 : mise à niveau de Thrift vers la version 0.13.0

Type	Description
Rétroportage	HIVE-24378 : Les espaces de début et de fin ne sont pas supprimés avant la conversion décimale
Rétroportage	HIVE-21341 : Valeurs par défaut raisonnables : hive.server2.idle.operation.timeout et hive.server2.idle.session.timeout sont trop élevés
Rétroportage	HIVE-22465 : Ajouter la configuration SSL dans TezConfigurationFactory
Rétroportage	HIVE-24710 : Optimisation de l'itération PTF pour le nombre (*) afin de réduire les coûts du processeur et des E/S
Rétroportage	HIVE-15406 : Envisagez de vectoriser la nouvelle fonction « trunc »
Rétroportage	HIVE-21541 : correction des en-têtes asf manquants dans HIVE-15406
Rétroportage	HIVE-24808 : Dates analysées dans le cache
Rétroportage	HIVE-24746 : PTF : TimestampValueBoundaryScanner peut être optimisé lors du calcul de la portée
Rétroportage	HIVE-25059 : L'événement Alter est converti pour être renommé lors de la réplication
Rétroportage	HIVE-25142 : Le remaniement dans la table de hachage rapide de la jointure cartographique entraîne la corruption des clés volumineuses
Rétroportage	HIVE-23756 : Ajout de contraintes supplémentaires au fichier package.jdo

Type	Description
Rétroportage	HIVE-25150 : Les caractères de tabulation ne sont pas supprimés avant la conversion décimale, comme le caractère d'espace qui est corrigé dans le cadre de HIVE-24378
Rétroportage	HIVE-25093 : date_format () UDF renvoie une sortie uniquement dans le fuseau horaire UTC
Rétroportage	HIVE-25268 : date_format udf renvoie des résultats erronés pour les dates antérieures à 1900 si le fuseau horaire local est autre que UTC
Rétroportage	HIVE-25338 : AIOBE dans une configuration UDF si l'entrée est vide
Rétroportage	HIVE-22400 : la minute UDF avec le temps renvoie la valeur NULL
Rétroportage	HIVE-25058 : PTF : TimestampValueBoun daryScanner peut être optimisé lors du calcul de la plage pt2 - isDistanceGreater
Rétroportage	HIVE-25449 : datediff () donne un mauvais résultat lorsqu'il est exécuté dans une tâche tez avec un fuseau horaire autre que l'UTC
Rétroportage	HIVE-23688 : Vectorisation : IndexArra yOutOfBoundsException pour une colonne de type carte qui inclut une valeur nulle
Rétroportage	HIVE-22247 : Hive se HFile OutputFormat lance FileNotFoundException lorsque la sortie de tâche de la partition est vide

Type	Description
Rétroportage	HIVE-25570 : Hive doit envoyer le chemin URL complet pour l'autorisation de la commande insert overwrite location
Rétroportage	HIVE-22903 : La vectorisation row_number () réinitialise le numéro de ligne après un lot en cas d'expression constante dans la clause de partition
Rétroportage	HIVE-25549 : Résultats erronés pour la fonction de fenêtre avec une expression dans la clause PARTITION BY ou ORDER BY
Rétroportage	HIVE-25579 : LOAD overwrite ajoute plutôt que de remplacer
Rétroportage	HIVE-25659 : Les requêtes SQL directes de Metastore avec IN/ (NOT IN) doivent être divisées en fonction du nombre maximum de paramètres autorisés par SQL DB
Rétroportage	HIVE-20502 : Corrigez le NPE lors de l'exécuti on de skewjoin_mapjoin10.q lorsque les statistiques des colonnes sont utilisées.
Rétroportage	HIVE-25765 : la propriété skip.header.line.c ount ignore les lignes de chaque bloc lorsque la taille du fichier est plus grande FetchOperator
Bogue	Corrigez NPE lors de l'insertion dans certains scénarios lorsque `hive.stats.column. autogather` et `hive.groupby.skewi ndata` sont tous les deux activés
Bogue	Corriger NPE lorsque la valeur `mapred.ta sktracker.expiry.interval` n'est pas définie

Amazon EMR 6.9.0 - Notes de mise à jour de Hive

Amazon EMR 6.9.0 - Changements concernant Hive

Type	Description
Upgrade	Mise à niveau de Jetty vers 9.4.48.v20220622
Upgrade	Prise en charge pour Hadoop 3.3.3.
Fonctionnalité	Intégration d'Amazon EMR Hive à Lake Formation pour les charges de travail interactives à l'aide de l'API GCSC.
Fonctionnalité	Intégration de Hive d'Amazon EMR avec Iceberg.
Amélioration	Activez le protocole SSL en HiveServer 2 lorsque le chiffrement en transit est activé à l'aide des configurations de sécurité Amazon EMR.
Amélioration	Activez le commit optimisé Hive EMRFS pour Amazon S3 par défaut. Pour plus d'informations, voirActivation du validateur optimisé Hive EMRFS S3.
Amélioration	Ajoutez `HiveHBaseTableInputFormatV2` qui hérite uniquement de la version mappée de InputFormat pour corriger SPARK-34210. Définissez `hive.hbase.inputformat.v2` sur `true` pour l'utiliser.
Amélioration	Attendez que TeZam se lance en arrière-plan avec hive.cli.tez.session.async au lieu de l'arrêter et d'en lancer un nouveau immédiatement. Utilisez `hive.emr.cli.tez.session.open.timeout` pour définir ce délai en secondes.

Type	Description
Amélioration	Ajoutez l'option hive.conf.restricted.list.append pour ajouter les configurations séparées par des virgules à la liste des configurations restreintes existante `hive.conf.restricted.list` .
Amélioration	Message d'erreur plus clair lorsque la requête Hive échoue car l'emplacement n'est pas défini pour la base de données.
Rétroportage	HIVE-24484 : mise à niveau de Hadoop vers la version 3.3.1 et de Tez vers la version 0.10.2
Rétroportage	HIVE-22398 : Supprimez la gestion des files d'attente YARN via. ShimLoader
Rétroportage	HIVE-23190 : LLAP : modifier IndexCache pour transmettre l'objet du système de fichiers à. TezSpillRecord
Rétroportage	HIVE-22185 : HADOOP-15832 posera des problèmes lors des tests utilisant des clusters. MiniYarn
Rétroportage	HIVE-21670 : Remplacement par `mockito-all` dépendance `mockito-core` .
Rétroportage	HIVE-24542 : Préparez la goyave pour les améliorations.
Rétroportage	HIVE-23751 : QTest : Remplacez la `#mkdirs()` méthode pour l'aligner après HADOOP-16582`ProxyFileSystem` .
Rétroportage	HIVE-21603 : Préparation à Java 11 : mise à jour de la version PowerMock.

Type	Description
Rétroportage	HIVE-24083 : `hcatalog` erreur dans Hadoop 3.3.0 : type d'authentification requis.
Rétroportage	HIVE-24282 : Afficher les colonnes ne doit pas trier les colonnes de sortie sauf mention explicite.
Rétroportage	HIVE-20656 : Valeurs par défaut sensibles : les configurations de mémoire d'agrégation de cartes sont trop agressives.
Rétroportage	HIVE-25443 : Arrow SerDe ne peut pas sérialiser/désérialiser les types de données complexes lorsqu'il y a plus de 1024 valeurs
Rétroportage	HIVE-19792 : Mise à niveau de orc vers la version 1.5.2 et activez les tests d'évolution du schéma decimal_64.
Rétroportage	HIVE-20437 : Gérer l'évolution du schéma à partir de nombres flottants, doubles et décimaux.
Rétroportage	HIVE-21987 : Hive ne parvient pas à lire Parquet int32 annoté en décimal.
Rétroportage	HIVE-20038 : Les requêtes de mise à jour sur des tables non segmentées et partitionnées génèrent un NPE.

Amazon EMR 6.9.0 - Problèmes connus liés à Hive

- Avec Amazon EMR 6.6.0 à 6.9.x, les requêtes INSERT avec partition dynamique et clause ORDER BY ou SORT BY auront toujours deux réducteurs. Ce problème est dû à la modification d'OSS HIVE-20703, qui place l'optimisation des partitions dynamiques de tri dans le cadre d'une décision basée sur les coûts. Si votre charge de travail ne nécessite

pas le tri des partitions dynamiques, nous vous recommandons de définir la propriété `hive.optimize.sort.dynamic.partition.threshold` sur `-1` pour désactiver la nouvelle fonctionnalité et obtenir le nombre de réducteurs correctement calculé. Ce problème est résolu dans OSS Hive dans le cadre de HIVE-22269 et dans Amazon EMR 6.10.0.

Amazon EMR 6.8.0 - Notes de mise à jour de Hive

Amazon EMR 6.8.0 - Changements concernant Hive

Type	Description
Amélioration	Réduisez les appels au système de fichiers dans la commande msck. Améliorations des performances (environ 15 à 20 fois sur plus de 10 000 partitions)
Rétroportage	HIVE-20678 : Hive HBase TableOutputFormat doit être implémenté pour garantir la compatibi lité HiveOutputFormat
Rétroportage	HIVE-21040 : msck fait une liste de fichiers inutiles au dernier niveau de l'arborescence des répertoires
Rétroportage	HIVE-21460 : Le chargement des données suivi d'une requête select * entraîne des résultats incorrects
Rétroportage	HIVE-21660 : Mauvais résultat lors de l'utilisa tion de tout l'union et de la vue ultérieure avec explode
Rétroportage	HIVE-22505 : ClassCastException causé par une mauvaise sélection d'opérateur vectorisé
Rétroportage	HIVE-22513 : La propagation constante de la colonne coulée lors des opérations de filtrage peut entraîner des résultats incorrects

Type	Description
Rétroportage	HIVE-23435 : Résultat complet de la jointure externe : lignes manquantes
Rétroportage	HIVE-24209 : Conversion incorrecte des arguments de recherche pour l'opération NOT BETWEEN lorsque la vectorisation est activée
Rétroportage	HIVE-24934 : aucune VectorizedExpressi ons annotation n'est nécessaire dans Generic Check UDFSQCount
Rétroportage	HIVE-25278 : HiveProjectJoinTransposeRule peut effectuer des transformations non valides avec des expressions de fenêtrage
Rétroportage	HIVE-25505 : Résultats incorrects avec en-tête. skip.header.line.count si la première ligne est vide
Rétroportage	HIVE-26080 : Mise à niveau d'accumulo-core vers la version 1.10.1
Rétroportage	HIVE-26235 : La condition OR sur la colonne binaire renvoie un résultat vide
Bogue	Corriger les journaux d'avertissement de plusieurs liaisons SLF4 J dans stderr lors du lancement
Bogue	Corrigez l'échec de la requête SHOW TABLE EXTENDED avec une erreur Wrong FS lorsque la partition et la table se trouvent sur des systèmes de fichiers différents.

Amazon EMR 6.8.0 - Problèmes connus liés à Hive

- Avec Amazon EMR 6.6.0 à 6.9.x, les requêtes INSERT avec partition dynamique et clause ORDER BY ou SORT BY auront toujours deux réducteurs. Ce problème est dû à la modification d'OSS HIVE-20703, qui place l'optimisation des partitions dynamiques de tri dans le cadre d'une décision basée sur les coûts. Si votre charge de travail ne nécessite pas le tri des partitions dynamiques, nous vous recommandons de définir la propriété `hive.optimize.sort.dynamic.partition.threshold` sur `-1` pour désactiver la nouvelle fonctionnalité et obtenir le nombre de réducteurs correctement calculé. Ce problème est résolu dans OSS Hive dans le cadre de HIVE-22269 et dans Amazon EMR 6.10.0.

Amazon EMR 6.7.0 - Notes de mise à jour de Hive

Amazon EMR 6.7.0 - Changements concernant Hive

Type	Description
Fonctionnalité	Intégration d'Amazon EMR Hive avec. LakeFormation
Fonctionnalité	Journalisation d'audit supplémentaire pour le commit optimisé Hive EMRFS pour Amazon S3. Configuration Hive : `hive.blobstore.output-committer.logging`, par défaut : `false`
Fonctionnalité	Le répertoire cible supprimé lors de l'insertion est remplacé par un résultat de sélection vide dans une table/partition statique non partitionnée afin qu'il se comporte de la même manière que Hive 2.x. Configuration Hive : `hive.emr.iow.clean.target.dir`, par défaut : `false`
Bogue	Correction d'un échec de requête intermittent lors de l'utilisation du validateur Hive EMRFS

Type	Description
	Amazon S3 optimisé avec tri des compartiments de partition.
Upgrade	La version de Hive a été mise à niveau vers 3.1.3. Reportez-vous aux notes de mise à jour d'Apache Hive 3.1.3 pour plus de détails.
Upgrade	Parquet mis à niveau vers la version 1.12.2.
Rétroportage	HIVE-20065 : Metastore ne doit pas s'appuyer sur Jackson 1.x
Rétroportage	HIVE-20071 : Migrer vers Jackson 2.x et empêcher son utilisation
Rétroportage	HIVE-20607 : TxnHandler doit être utilisé PreparedStatement pour exécuter des requêtes SQL directes
Rétroportage	HIVE-20740 : Supprimer le verrou global dans la méthode .setConf ObjectStore
Rétroportage	HIVE-20961 : Suppression de l'implémentation de NVL
Rétroportage	HIVE-22059 : le fichier jar hive-exec ne contient pas la bibliothèque Jackson (fasterxml)
Rétroportage	HIVE-22351 : Correction d'une utilisation incorrecte des threads dans ObjectStore TestObjectStore
Rétroportage	HIVE-23534 : NPE dans RetryingMetaStoreClient #invoke lors d'une capture sans message MetaException
Rétroportage	HIVE-24048 : Harmoniser les composants Jackson avec la version 2.10.latest - Hive

Type	Description
Rétroportage	HIVE-24768 : Utiliser jackson-bom partout pour le remplacement de version
Rétroportage	HIVE-24816 : Mise à niveau de Jackson vers la version 2.10.5.1 ou 2.11.0+ en raison de CVE-2020-25649
Rétroportage	HIVE-25971 : L'arrêt de la tâche Tez est retardé car le pool de threads mis en cache n'est pas fermé
Rétroportage	HIVE-26036 : NPE causé par get () in MTable ObjectStore

Amazon EMR 6.7.0 - Problèmes connus liés à Hive

- Les requêtes avec des fonctions de fenêtrage sur la même colonne que la jointure peuvent conduire à des transformations invalides comme rapporté dans HIVE-25278 et causer des résultats incorrects ou des échecs de requête. Une solution de contournement serait de désactiver le CBO au niveau de la requête pour de telles requêtes. Le correctif sera disponible dans une version Amazon EMR ultérieure à la version 6.7.0. Pour plus d'informations, contactez AWS le support.
- Avec Amazon EMR 6.6.0 à 6.9.x, les requêtes INSERT avec partition dynamique et clause ORDER BY ou SORT BY auront toujours deux réducteurs. Ce problème est dû à la modification d'OSS HIVE-20703, qui place l'optimisation des partitions dynamiques de tri dans le cadre d'une décision basée sur les coûts. Si votre charge de travail ne nécessite pas le tri des partitions dynamiques, nous vous recommandons de définir la propriété `hive.optimize.sort.dynamic.partition.threshold` sur `-1` pour désactiver la nouvelle fonctionnalité et obtenir le nombre de réducteurs correctement calculé. Ce problème est résolu dans OSS Hive dans le cadre de HIVE-22269 et dans Amazon EMR 6.10.0.

Amazon EMR 6.6.0 - Notes de mise à jour de Hive

Amazon EMR 6.6.0 - Changements concernant Hive

Type	Description
Upgrade	Mettez à niveau de Parquet vers la version 1.12.1.
Upgrade	Mettre à niveau la version Jetty Jars vers 9.4.43.v20210629
Bogue	Correction d'un problème qui provoquait l'installation de Hive sur tous les nœuds de tâche/cœur lorsque LLAP était activé sur un cluster Hive.
Rétroportage	HIVE-25942 : Mise à niveau de commons-io vers la version 2.8.0 en raison de CVE-2021-29425
Rétroportage	HIVE-25726 : amélioration de la vélocité à 2.3 en raison de CVE-2020-13936
Rétroportage	HIVE-25680 : Autorisez l'API HiveMetastore du serveur #get_table_meta à utiliser l'un des modèles d'autorisation. HiveMetastore
Rétroportage	HIVE-25554 : Mise à niveau de la version Arrow vers 0.15
Rétroportage	HIVE-25242 : La requête s'exécute extrêmement lentement avec vectorized.adaptor = chosen
Rétroportage	HIVE-25085 : MetaStore Les clients ne sont plus partagés entre les sessions.
Rétroportage	HIVE-24827 : La requête d'agrégation Hive renvoie des résultats incorrects pour les fichiers autres que du texte.

Type	Description
Rétroportage	HIVE-24683 : getFileId Hadoop23Shims est sujet à NPE pour les chemins inexistants
Rétroportage	HIVE-24656 : Le CBO échoue pour les requêtes dont la valeur est nulle sur les types de carte et de tableau
Rétroportage	HIVE-24556 : Optimisé DefaultGraphWalker pour les cas sans petit-enfant
Rétroportage	HIVE-24408 : Mise à niveau de Parquet vers la version 1.11.1
Rétroportage	HIVE-24391 : Corriger les échecs TestOrcFile FIX dans la branche 3.1
Rétroportage	HIVE-24362 : Le traitement des arbres AST n'est pas optimal pour les arbres comportant un grand nombre de nœuds
Rétroportage	HIVE-24316 : Mise à niveau d'ORC de la version 1.5.6 à la version 1.5.8 dans la branche 3.1
Rétroportage	HIVE-24307 : Beeline avec le fichier de propriétés et le paramètre -e échoue
Rétroportage	HIVE-24245 : PTF vectorisé avec nombre et partition distincte produisant des résultats incorrects.
Rétroportage	HIVE-24224 : Correction du saut d'en-tête/ pied de page pour Hive on Tez sur un fichier compressé
Rétroportage	HIVE-24157 : Le mode strict échouera sur CAST timestamp ↔ numeric

Type	Description
Rétroportage	HIVE-24113 : NPE dans le générique UDFTo UnixTimeStamp
Rétroportage	HIVE-23987 : Mise à niveau de arrow vers la version 0.11.0
Rétroportage	HIVE-23972 : Ajouter un identifiant de client externe au client externe LLAP
Rétroportage	HIVE-23806 : Éviter d'effacer les états des statistiques des colonnes dans toutes les partitions au cas où le schéma serait étendu. Cela améliore l'exécution de l'instruction alter table add columns.
Rétroportage	HIVE-23779 : BasicStatsTask Les informations ne sont pas imprimées dans la console Beeline
Rétroportage	HIVE-23306 : La commande RESET ne fonctionne pas s'il existe une configuration définie par System.GetProperty
Rétroportage	HIVE-23164 : Le serveur n'est pas correctem ent arrêté à cause de threads autres que les démons
Rétroportage	HIVE-22967 : Support hive.reloadable.au x.jars.path pour Hive on Tez
Rétroportage	HIVE-22934 : Le serveur Hive compte les journaux interactifs sur le flux d'erreurs
Rétroportage	HIVE-22901 : La substitution de variables peut entraîner un OOM sur les références circulaires

Type	Description
Rétroportage	HIVE-22769 : Résultats de requête incorrects et échec de requête lors de la génération fractionnée pour les fichiers texte compressés
Rétroportage	HIVE-22716 : La lecture est interrompue ByteBuffer ParquetFooterInputFromCache
Rétroportage	HIVE-22648 : Mise à niveau de Parquet vers la version 1.11.0
Rétroportage	HIVE-22640 : Decimal64 ColumnVector : ClassCastException lorsque le type de colonne de partition est Decimal
Rétroportage	HIVE-22621 : cas de test instable : .testSigning TestLlapSignerImpl
Rétroportage	HIVE-22533 : Corrige les vulnérabilités possibles de l'interface utilisateur Web du démon LLAP
Rétroportage	HIVE-22532 : Le PTFPPD peut pousser la limite de manière incorrecte via Rank/Function DenseRank
Rétroportage	HIVE-22514 : HiveProtoLoggingHook peut consommer beaucoup de mémoire
Rétroportage	HIVE-22476 : La fonction Hive datediff a fourni des résultats incohérents lorsque hive.fetch.task.conversion est défini sur none
Rétroportage	HIVE-22429 : Les tables clusterisées migrées à l'aide de bucketing_version 1 sur Hive 3 utilisent bucketing_version 2 pour les insertions

Type	Description
Rétroportage	HIVE-22412 : StatsUtils lance un NPE lorsque vous expliquez
Rétroportage	HIVE-22360 : MultiDelimitSerDe renvoie des résultats erronés dans la dernière colonne lorsque le fichier chargé comporte plus de colonnes que celles du schéma de table
Rétroportage	HIVE-22332 : Hive doit garantir des paramètres d'évolution de schéma valides depuis ORC-540
Rétroportage	HIVE-22331 : unix_timestamp sans argument renvoie l'horodatage en millisecondes au lieu de secondes
Rétroportage	HIVE-2275 :. OperationManager queryIdOperationne nettoie pas correctement plusieurs QueryID
Rétroportage	HIVE-22273 : La vérification d'accès échoue lorsqu'un répertoire temporaire est supprimé
Rétroportage	HIVE-22270 : Mise à niveau de commons-io vers la version 2.6
Rétroportage	HIVE-22241 : Implémenter l'UDF pour interpréter la date/l'horodatage à l'aide de sa représentation interne et du calendrier hybride grégorien-julien
Rétroportage	HIVE-22241 : Implémenter l'UDF pour interpréter la date/l'horodatage à l'aide de sa représentation interne et de son hybride grégorien-julien
Rétroportage	HIVE-22232 : NPE lorsque hive.order.columnalignment est défini sur false

Type	Description
Rétroportage	HIVE-22231 : Échec de l'écriture d'une requête Hive de grande taille via Knox en raison d'un tube cassé
Rétroportage	HIVE-22221 : Client externe Llap - Nécessité de réduire le #getSplits LlapBaseInputFormat
Rétroportage	HIVE-22208 : Le nom de colonne avec le mot clé réservé n'est pas échappé lorsque la requête incluant une jointure sur une table avec une colonne de masque est réécrite
Rétroportage	HIVE-22197 : Common Merge Join lance une exception de conversion de classe.
Rétroportage	HIVE-22170 : from_unixtime et unix_timestamp doivent utiliser le fuseau horaire de la session utilisateur
Rétroportage	HIVE-22169 : Tez : SplitGenerator essaie de rechercher des fichiers de plan qui n'existeront pas pour Tez
Rétroportage	HIVE-22168 : Supprimer la journalisation très coûteuse du hotpath du cache LLAP
Rétroportage	HIVE-22161 : UDF : synchronise sur org.apache.hadoop.hive.ql.udf. FunctionRegistry UDFType classe
Rétroportage	HIVE-22120 : Correction de résultats erronés ou d'une ArrayOutOfBound exception dans les jointures cartographiques externes gauches liées à des conditions limites spécifiques

Type	Description
Rétroportage	HIVE-22115 : Empêcher la création d'un appendice de routage de requêtes si la propriété est définie sur false
Rétroportage	HIVE-22113 : Empêcher l'arrêt du LLAP en cas de problème AMReporter RuntimeException
Rétroportage	HIVE-22106 : Supprimer la synchronisation des requêtes croisées pour la partition-eval
Rétroportage	HIVE-22099 : Plusieurs dates liées à des dates ne UDFs peuvent pas gérer correctement les dates juliennes depuis HIVE-20007
Rétroportage	HIVE-22037 : HS2 devrait être enregistré lors de l'arrêt dû à l'OOM
Rétroportage	HIVE-21976 : Le décalage doit être nul au lieu de zéro dans Calcite HiveSortLimit
Rétroportage	HIVE-21924 : Diviser les fichiers texte même s'il existe un en-tête/pied de page
Rétroportage	HIVE-21913 : Generic UDTFGet Splits doit gérer les noms d'utilisateur de la même manière que LLAP
Rétroportage	HIVE-21905 : Amélioration des génériques dans la classe FetchOperator
Rétroportage	HIVE-21902 : interface utilisateur HiveServe r 2 : besoins de l'en-tête de réponse Jetty X-Frame-Options
Rétroportage	HIVE-21888 : Définir hive.parquet.times tamp.skip.conversion par défaut sur true

Type	Description
Rétroportage	HIVE-21868 : Vectoriser CAST... FORMAT
Rétroportage	HIVE-21864 : #closeAll LlapBaseInputFormat
Rétroportage	HIVE-21863 : Améliorer le casting de type vectorisateur pour l'expression WHEN
Rétroportage	HIVE-21862 : ORC ppd produit un résultat erroné avec l'horodatage
Rétroportage	HIVE-21846 : Créez un thread dans TeZam qui récupère périodiquement des métriques LlapDaemon
Rétroportage	HIVE-21837 : MapJoin lance une exception lorsque la colonne sélectionnée contient des valeurs complètement nulles
Rétroportage	HIVE-21834 : Éviter les appels inutiles pour simplifier les conditions de filtrage
Rétroportage	HIVE-21832 : Nouvelles mesures pour obtenir le temps moyen queue/serving/response
Rétroportage	HIVE-21827 : Plusieurs appels dans la méthode SemanticAnalyzer Do Not Go Through getTableObject ByName
Rétroportage	HIVE-21822 : Exposez les LlapDaemon métriques via une nouvelle méthode d'API
Rétroportage	HIVE-21818 : CBO : la copie contient du trafic de métastore TableRelOptHiveTable
Rétroportage	HIVE-21815 : Les statistiques du fichier ORC sont analysées deux fois

Type	Description
Rétroportage	HIVE-21805 : HiveServer 2 : Utilisez le fast ShutdownHookManager APIs
Rétroportage	HIVE-21799 : NullPointerException activé DynamicPartitionPruningOptimization, lorsque la clé de jointure se trouve sur la colonne d'agrégation
Rétroportage	HIVE-21794 : Ajouter des paramètres de vue matérialisée à sqlStdAuth SafeVarNameRegexes
Rétroportage	HIVE-21768 : JDBC : supprime le préfixe d'union par défaut pour les requêtes UNION non incluses
Rétroportage	HIVE-21746 : ArrayIndexOutOfBoundsException lors d'une jointure par hachage partitionnée dynamiquement, avec CBO désactivé
Rétroportage	HIVE-21717 : Impossible de renommer le répertoire dans la tâche de déplacement.
Rétroportage	HIVE-21685 : Mauvaise simplification dans une requête contenant plusieurs clauses IN
Rétroportage	HIVE-21681 : La description formatée montre des informations incorrectes pour plusieurs clés primaires
Rétroportage	HIVE-21651 : Déplace le serveur protobuf dans hive-exec.
Rétroportage	HIVE-21619 : Afficher le type d'horodatage sans précision dans SQL Extended

Type	Description
Rétroportage	HIVE-21592 : OptimizedSql n'apparaît pas lorsque l'expression contient CONCAT
Rétroportage	HIVE-21576 : Présentation de CAST... FORMAT et de la liste limitée des formats date/heure SQL:2016
Rétroportage	HIVE-21573 : Le transport binaire doit ignorer le principal si l'authentification est définie sur delegationToken
Rétroportage	HIVE-21550 : les TestObjectStore tests sont flous - Impossible d'obtenir un verrou dans le délai demandé
Rétroportage	HIVE-21544 : La propagation constante corrompt coalesce/case/when les expressions lors du pliage
Rétroportage	HIVE-21539 : GroupBy + la clause where sur la même colonne entraîne une réécriture incorrecte de la requête
Rétroportage	HIVE-21538 : Beeline : source du mot de passe même si le lecteur de console n'est pas passé au paramètre de connexion
Rétroportage	HIVE-21509 : LLAP peut mettre en cache des vecteurs de colonnes corrompus et renvoyer un résultat de requête erroné
Rétroportage	HIVE-21499 : ne doit pas supprimer la fonction du registre si la commande de création a échoué avec AlreadyExistsException

Type	Description
Rétroportage	HIVE-21496 : Le dimensionnement automatiq ue d'une mémoire tampon non ordonnée peut déborder
Rétroportage	HIVE-21468 : distinction majuscules/minuscu les dans les noms d'identifiant pour le gestionnaire de stockage JDBC
Rétroportage	HIVE-21467 : Supprimer les importations JUnit.framework.assert obsolètes
Rétroportage	HIVE-21435 : LlapBaseInputFormat devrait obtenir le numéro de tâche dans la configura tion TASK_ATTEMPT_ID s'il est présent, lors de la construction SubmitWorkRequestProto
Rétroportage	HIVE-21389 : La distribution Hive manque javax.ws.rs-api.jar après HIVE-21247
Rétroportage	HIVE-21385 : Autoriser la désactivation du transfert des calculs non séparables vers les sources JDBC
Rétroportage	HIVE-21383 : Gestionnaire de stockage JDBC : utilisez le catalogue et le schéma pour récupérer les tables si cela est spécifié
Rétroportage	HIVE-21382 : Optimisation de la réduction des groupes par clés - les clés ne sont pas réduites dans la requête 23
Rétroportage	HIVE-21362 : Ajoutez un format d'entrée et serde pour lire à partir de fichiers protobuf.
Rétroportage	HIVE-21340 : CBO : élaguez les colonnes non essentielles alimentant un SemiJoin

Type	Description
Rétroportage	HIVE-21332 : Purger les tampons non verrouillés au lieu des tampons verrouillés
Rétroportage	HIVE-21329 : Taille de la mémoire tampon de sortie non ordonnée de l'exécution Tez personnalisée en fonction du pipeline de l'opérateur
Rétroportage	HIVE-21295 : StorageHandler doit convertir la date en chaîne en utilisant la convention Hive
Rétroportage	HIVE-21294 : Vectorisation : Shuffle à 1 réducteur peut ignorer les fonctions de hachage d'objets
Rétroportage	HIVE-21255 : Supprimer dans QueryConditionBuilder JdbcStorageHandler
Rétroportage	HIVE-21253 : Support dans JDBC DB2 StorageHandler
Rétroportage	HIVE-21232 : LLAP : ajout d'un fournisseur d'affinité partagée compatible avec Cache-miss
Rétroportage	HIVE-21214 : MoveTask : Utilisez AttentID au lieu de la taille du fichier pour la déduplication des fichiers compareTempOr DuplicateFiles
Rétroportage	HIVE-21184 : Ajouter, expliquer et expliquer un plan CBO formaté avec des informations sur les coûts
Rétroportage	HIVE-21182 : Ignorer la configuration de Hive Scratch Dir lors de la planification
Rétroportage	HIVE-21171 : Ignorer la création de scratch pour Tea si le RPC est activé

Type	Description
Rétroportage	HIVE-21126 : Autoriser les requêtes au niveau de la session dans #getSplit LlapBaseI nputFormat
Rétroportage	HIVE-21107 : Erreur « Impossible de trouver le champ » lors d'une jointure par hachage partitionnée dynamiquement
Rétroportage	HIVE-21061 : La requête CTAS échoue avec une source vide IllegalStateException
Rétroportage	HIVE-21041 : NPE, lors de l'obtention d'un schéma ParseException à partir d'un plan logique
Rétroportage	HIVE-21013 : Impossible de trouver JdbcStora geHandler la colonne de partition dans Oracle
Rétroportage	HIVE-21006 : Étendre SharedWorkOptimize r pour supprimer les semi-jointures lorsqu'il existe une opportunité de réutilisation
Rétroportage	HIVE-20992 : Diviser la configuration hive.meta store.dbaccess.ssl.properties en configurations plus significatives
Rétroportage	HIVE-20989 : JDBC - Le journal GetOperat ionStatus + peut bloquer la progression des requêtes via le mode veille
Rétroportage	HIVE-20988 : Résultats erronés pour les requêtes de groupe par avec une clé primaire sur plusieurs colonnes

Type	Description
Rétroportage	HIVE-20985 : Si les entrées de l'opérateur sélectionné sont des colonnes temporaires, la vectorisation peut réutiliser certaines d'entre elles comme sortie
Rétroportage	HIVE-20978 : « hive.jdbc.* » doit être ajouté à sqlStdAuth SafeVarNameRegexes
Rétroportage	HIVE-20953 : Supprime une fonction du registre des fonctions lorsqu'elle ne peut pas être ajoutée au métastore lors de sa création.
Rétroportage	HIVE-20952 : Nettoyage d'un fichier .java VectorizationContext
Rétroportage	HIVE-20951 : LLAP : réglez toujours Xms à 50 %
Rétroportage	HIVE-20949 : Améliorer l'estimation de la cardinalité PKFK dans le cadre de la planifica tion physique
Rétroportage	HIVE-20944 : Les statistiques ne sont pas validées lors de la compilation des requêtes
Rétroportage	HIVE-20940 : Réduire les cas dans lesquels la résolution de type de Calcite est plus stricte que celle de Hive.
Rétroportage	HIVE-20937 : La requête JDBC de Postgres échoue avec le message « LIMIT ne doit pas être négatif »
Rétroportage	HIVE-20926 : L'indicateur de réduction des semi-jointures échoue lorsque le nombre d'entrées du filtre Bloom est élevé ou lorsqu'il n'y a pas de statistiques

Type	Description
Rétroportage	HIVE-20920 : Utiliser les contraintes SQL pour améliorer l'algorithme de réorganisation des jointures
Rétroportage	HIVE-20918 : Indicateur pour activer/désactiver le transfert du calcul depuis Calcite vers une connexion JDBC
Rétroportage	HIVE-20915 : Rendre l'optimisation des partitions de tri dynamique accessible à HoS et MR
Rétroportage	HIVE-20910 : L'insertion dans une table à compartiments échoue en raison de l'optimis ation dynamique du tri des partitions
Rétroportage	HIVE-20899 : L'URI Keytab pour le service LLAP YARN est restrictif pour prendre en charge uniquement HDFS
Rétroportage	HIVE-20898 : Pour les fonctions liées au temps, les arguments ne peuvent pas être convertis en un type non nullable
Rétroportage	HIVE-20881 : La propagation constante simplifie à outrance les projections
Rétroportage	HIVE-20880 : Mettre à jour la valeur par défaut pour hive.stats.filter.in.min.ratio
Rétroportage	HIVE-20873 : Utilisez le hachage Murmur pour réduire les collisions de hachage VectorHas hKeyWrapperTwoLong

Type	Description
Rétroportage	HIVE-20868 : SMB Join échoue par intermittence lorsqu'un enfant se connecte TezDummyOperator getFinalOp MapRecord Processor
Rétroportage	HIVE-20853 : Expose ShuffleHandler .register Dag dans l'API du démon llap
Rétroportage	HIVE-20850 : Transférer les projections aux tableaux de dimensions si possible
Rétroportage	HIVE-20842 : Correction de la logique introduite dans HIVE-20660 pour estimer les statistiques du groupe par
Rétroportage	HIVE-20839 : Erreur « Impossible de trouver le champ » lors d'une jointure par hachage partitionnée dynamiquement
Rétroportage	HIVE-20835 : L'interaction entre les contraint es et la réécriture de MV peut créer une boucle dans Calcite Planner
Rétroportage	HIVE-20834 : QueryResultCache entrées Hive conservant la référence à la requête mise en cache SemanticAnalyzer
Rétroportage	HIVE-20830 : échec de l'assertion JdbcStora geHandler de requête de plage dans certains cas
Rétroportage	HIVE-20829 : JdbcStorageHandler Range Split lance un NPE
Rétroportage	HIVE-20827 : Résultats incohérents pour les tableaux vides

Type	Description
Rétroportage	HIVE-20826 : amélioration de la HiveSemiJ oin règle pour convertir la jointure et le groupe situés sur le côté gauche en semi-jointure gauche
Rétroportage	HIVE-20821 : Réécrire SUM0 en combinaison SUM+ COALESCE
Rétroportage	HIVE-20815 : JdbcRecordReader .next ne doit pas manger d'exception
Rétroportage	HIVE-20813 : udf to_epoch_milli doit également prendre en charge l'horodatage sans fuseau horaire.
Rétroportage	HIVE-20804 : Améliorations supplémentaires apportées au regroupement par optimisation avec contraintes
Rétroportage	HIVE-20792 : L'insertion d'un horodatage avec des zones tronque les données
Rétroportage	HIVE-20788 : La réduction SJ étendue peut revenir en arrière de manière incorrecte lors de la création de filtres
Rétroportage	HIVE-20778 : La réorganisation des jointures peut ne pas être déclenchée si toutes les jointures du plan sont créées par une logique de décorrélation
Rétroportage	HIVE-20772 : Enregistre les compteurs de CPU par tâche dans LLAP
Rétroportage	HIVE-20768 : Ajout d'une fenêtre Tumbling UDF

Type	Description
Rétroportage	HIVE-20767 : Plusieurs projets entre opérateurs de jointure peuvent affecter la réorganisation des jointures en utilisant des contraintes
Rétroportage	HIVE-20762 : L'intervalle de nettoyage de NOTIFICATION_LOG est codé en dur à 60 secondes et est trop petit
Rétroportage	HIVE-20761 : La sélection pour la mise à jour sur la table notification_sequence a un intervalle de tentatives et un nombre de tentatives trop faible
Rétroportage	HIVE-20751 : Mise à niveau de arrow vers la version 0.10.0
Rétroportage	HIVE-20746 : HiveProtoHookLogger ne ferme pas le fichier en fin de journée.
Rétroportage	HIVE-20744 : Utiliser les contraintes SQL pour améliorer l'algorithme de réorganisation des jointures
Rétroportage	HIVE-20740 : Supprime le verrou global dans ObjectStore la méthode .setConf. Ce cherrypick rétroporte le HIVE-20740 destiné à Hive 3.2 et 4.x vers 3.1.x
Rétroportage	HIVE-20734 : Beeline : lorsque beeline-site.xml est activé et que la CLI hive redirige vers beeline, elle doit utiliser le nom d'utilisateur/mot de passe fictif du système au lieu de demander un
Rétroportage	HIVE-20731 : le fichier keystore doit être autorisé JdbcStorageHandler

Type	Description
Rétroportage	HIVE-20720 : Ajouter une option de colonne de partition au gestionnaire JDBC
Rétroportage	HIVE-20719 : L'instruction SELECT échoue après la mise à jour avec l'optimisation et la vectorisation de hive.optimize.sort.dynamic. partition activées
Rétroportage	HIVE-20718 : Ajout d'un pilote Perf CLI avec des contraintes
Rétroportage	HIVE-20716 : Définir la valeur par défaut de hive.cbo.stats.correlated.multi.key.joins sur true
Rétroportage	HIVE-20712 : HivePointLookupOptimizer devrait extraire les cas profonds
Rétroportage	HIVE-20710 : Le pliage constant ne peut pas créer de constantes nulles sans types
Rétroportage	HIVE-20706 : external_jdbc_table2.q échoue par intermittence
Rétroportage	HIVE-20704 : extension HivePreFilteringRule pour prendre en charge d'autres fonctions
Rétroportage	HIVE-20703 : Placer l'optimisation des partition s de tri dynamique dans le cadre d'une décision basée sur les coûts
Rétroportage	HIVE-20702 : Tenir compte de la surcharge liée aux estimations tenant compte de la structure des données lors de la sélection de mapjoin
Rétroportage	HIVE-20692 : Activer le pliage des expressions NOT x IS (NOT) [TRUE|FALSE]

Type	Description
Rétroportage	HIVE-20691 : Corrigez org.apache.hadoop. hive.cli. TestMiniLlapCliDriver. testCliDriver[cttl]
Rétroportage	HIVE-20682 : L'exécution d'une requête asynchrone peut potentiellement échouer si le SessionHive partagé est fermé par le thread principal
Rétroportage	HIVE-20676 : HiveServer 2 : n' Privilege Synchronizerest pas défini sur le statut de daemon
Rétroportage	HIVE-20660 : L'estimation des groupes par statistiques pourrait être améliorée en limitant le nombre total de lignes à la table source
Rétroportage	HIVE-20652 : jointure JdbcStorageHandler push de deux sources de données différentes vers le pilote jdbc
Rétroportage	HIVE-20651 : JdbcStorageHandler le mot de passe doit être crypté
Rétroportage	HIVE-20649 : Gestionnaire de mémoire compatible LLAP pour les rédacteurs Orc
Rétroportage	HIVE-20648 : LLAP : le groupe de vecteurs par opérateur doit utiliser de la mémoire par exécuteur
Rétroportage	HIVE-20646 : La condition du filtre de partition n'est pas transmise à la requête Metastore si elle contient la valeur IS NOT NULL
Rétroportage	HIVE-20644 : Éviter d'exposer des informations sensibles via une exception Hive Runtime

Type	Description
Rétroportage	HIVE-20636 : Améliorer l'estimation du nombre de valeurs nulles après une jointure externe
Rétroportage	HIVE-20632 : La requête avec get_splits UDF échoue si une vue matérialisée est créée sur la table interrogée
Rétroportage	HIVE-20627 : Les requêtes asynchrones simultanées échouent par intermittence et provoquent une fuite de mémoire LockExcep tion
Rétroportage	HIVE-20623 : Travail partagé : étendre le partage des entrées du cache Map-Join dans LLAP
Rétroportage	HIVE-20619 : MultiDelimitSerDe Inclus dans 2 par défaut HiveServer
Rétroportage	HIVE-20618 : Lors de la jointure, la sélection BucketMapJoin peut être choisie pour les tables non groupées
Rétroportage	HIVE-20617 : Corrige le type de constantes dans les expressions IN pour qu'elles aient le type correct
Rétroportage	HIVE-20612 : Création d'un nouvel indicateur de corrélation multiclé de jointure pour CBO
Rétroportage	HIVE-20603 : Erreur « Wrong FS » lors de l'insertion dans la partition après avoir modifié l'emplacement de la table dans le système de fichiers

Type	Description
Rétroportage	HIVE-20601 : EnvironmentContext nul dans l'événement ALTER_PARTITION dans DbNotificationListener
Rétroportage	HIVE-20583 : Utiliser le nom d'hôte canonique uniquement pour l'authentification Kerberos dans HiveConnection
Rétroportage	HIVE-20582 : Rendre la journalisation de hflush dans Hive Proto configurable
Rétroportage	HIVE-20563 : Vectorisation : l'expression CASE WHEN échoue lorsque le type THEN/ELSE et le type de résultat sont différents
Rétroportage	HIVE-20558 : Remplacer la valeur par défaut de hive.hashtable.key.count.adjustment à 0,99
Rétroportage	HIVE-20552 : Obtenir le schéma plus rapidement LogicalPlan
Rétroportage	HIVE-20550 : Switch Web HCat pour utiliser beeline pour soumettre des requêtes Hive
Rétroportage	HIVE-20537 : Plusieurs colonnes joignent les estimations avec des colonnes non corrélées différentes dans CBO et Hive
Rétroportage	HIVE-20524 : La vérification de l'évolution du schéma est interrompue lors du passage de la version 2 à la version 3 de Hive pour ALTER TABLE VARCHAR à DECIMAL
Rétroportage	HIVE-20522 : HiveFilterSetOpTransposeRule peut générer une erreur d'assertion en raison de la nullité des champs

Type	Description
Rétroportage	HIVE-20521 : HS2 DOAS=true a un problème d'autorisation avec hadoop.tmp.dir, avec les systèmes de fichiers MR et S3A
Rétroportage	HIVE-20515 : Résultats de requête vides lors de l'utilisation du cache des résultats et du répertoire temporaire des requêtes, répertoir e du cache des résultats dans différents systèmes de fichiers
Rétroportage	HIVE-20508 : Hive ne prend pas en charge les noms d'utilisateur de type « user @realm »
Rétroportage	HIVE-20507 : Beeline : ajoute une commande utilitaire pour récupérer tous les URI du fichier beeline-site.xml
Rétroportage	HIVE-20505 : mise à niveau de org.openj dk.jmh:jmh-core vers la version 1.21
Rétroportage	HIVE-20503 : Utiliser des estimations tenant compte de la structure des données lors de la sélection de mapjoin
Rétroportage	HIVE-20498 : Type de date de support pour la collecte automatique des statistiques des colonnes
Rétroportage	HIVE-20496 : Vectorisation : PTF vectorisé IllegalStateException
Rétroportage	HIVE-20494 : Le générique UDFRestrict InformationSchema est cassé après HIVE-1944 0
Rétroportage	HIVE-20477 : n' OptimizedSql est pas affiché si l'expression contient INs

Type	Description
Rétroportage	HIVE-20467 : Autoriser SINON EXISTS/IF EXISTS in Resource plan creation/drop
Rétroportage	HIVE-20462 : « CREATE VIEW IF NOT EXISTS » échoue si la vue existe déjà
Rétroportage	HIVE-20455 : Extrait de journal provenant de security.authorization. PrivilegeSynchroniz er.exécuter
Rétroportage	HIVE-20439 : Utiliser la limite de mémoire exagérée lors de la sélection des jointures pour llap
Rétroportage	HIVE-20433 : La conversion implicite d'une chaîne en horodatage est lente
Rétroportage	HIVE-20432 : Réécrire BETWEEN en IN pour les types entiers pour l'estimation des statistiq ues
Rétroportage	HIVE-20423 : Définir NULLS LAST comme ordre nul par défaut
Rétroportage	HIVE-20418 : LLAP IO peut ne pas gérer les fichiers ORC dont l'index de ligne est correctem ent désactivé pour les requêtes sans colonne sélectionnée
Rétroportage	HIVE-20412 : Entrée NPE HiveMetaHook
Rétroportage	HIVE-20406 : La coalescence imbriquée donne des résultats incorrects
Rétroportage	HIVE-20399 : Le CTAS avec un emplacement de table personnalisé qui n'est pas entièrement qualifié échoue pour les tables MM

Type	Description
Rétroportage	HIVE-20393 : Réduction des semi-jointures : FordPP se comporte de manière incohérente markSemiJoin
Rétroportage	HIVE-20391 : HiveAggregateReduceFunctionsRule peut déduire un type de retour erroné lors de la décomposition de la fonction d'agrégation
Rétroportage	HIVE-20383 : Nom de file d'attente non valide et problèmes de synchronisation dans Hive Proto Events Hook.
Rétroportage	HIVE-20367 : Vectorisation : Support du streaming pour PTF AVG, MAX, MIN, SUM
Rétroportage	HIVE-20366 : Les estimations de statistiques de la requête TPC-DS 78 sont erronées pour son filtre nul
Rétroportage	HIVE-20364 : Mettre à jour la valeur par défaut pour hive.map.aggr.hash.min.reduction
Rétroportage	HIVE-20352 : Vectorisation : fonction de regroupement de supports
Rétroportage	HIVE-20347 : hive.optimize.sort.dynamic.partition devrait fonctionner avec des CTAS et des MV partitionnés
Rétroportage	HIVE-20345 : La base de données Drop peut se bloquer si les tables sont supprimées lors d'un autre appel
Rétroportage	HIVE-20343 : Hive 3 : le CTAS ne respecte pas les propriétés transactionnelles

Type	Description
Rétroportage	HIVE-20340 : Druid a besoin d'une valeur explicite CASTs allant de l'horodatage à la chaîne lorsque la sortie de la fonction d'horodatage est utilisée sous forme de chaîne
Rétroportage	HIVE-20339 : Vectorisation : suppression d'une restriction inutile empêchant la vectorisation de certains PTF avec RANK
Rétroportage	HIVE-20337 : CachedStore : getPartitionsByExpr ne remplit pas correctement la liste des partitions
Rétroportage	HIVE-20336 : Politiques de masquage et de filtrage pour les vues matérialisées
Rétroportage	HIVE-20326 : Créer des contraintes avec RELY par défaut au lieu de NO RELY
Rétroportage	HIVE-20321 : Vectorisation : réduction de la taille de la mémoire de 1 col à <1 VectorHashKeyWrapper CacheLine
Rétroportage	HIVE-20320 : Activer l'indicateur hive.optimize.remove.sq_count_check
Rétroportage	HIVE-20315 : Vectorisation : corriger davantage de problèmes de résultats nuls ou erronés et éviter les casts/conversions inutiles
Rétroportage	HIVE-20314 : Inclure l'élimination des partitions dans la réécriture des vues matérialisées
Rétroportage	HIVE-20312 : Autoriser les clients Arrow à utiliser les leurs avec BufferAllocator LlapOutputFormatService

Type	Description
Rétroportage	HIVE-20302 : LLAP : l'exécution non vectorisé e dans IO ignore les colonnes virtuelles, y compris ROW__ID
Rétroportage	HIVE-20300 : VectorFileSinkArrowOperator
Rétroportage	HIVE-20299 : course potentielle lors du test unitaire des signataires du LLAP
Rétroportage	HIVE-20296 : Améliorez HivePointLookupOpt imizerRule pour pouvoir extraire des données de contextes plus sophistiqués
Rétroportage	HIVE-20294 : Vectorisation : correction des problèmes de résultats nuls ou erronés dans COALESCE/ELT
Rétroportage	HIVE-20292 : Mauvais ordre des jointures dans la requête tpcds 93 avec une contrainte principale définie
Rétroportage	HIVE-20290 : Initialisation différée pour ne ArrowColumnarBatchSerDe pas allouer de tampons pendant GetSplits
Rétroportage	HIVE-20281 : SharedWorkOptimizer échec avec « le contenu du cache de l'opérateur et le plan réel diffèrent »
Rétroportage	HIVE-20277 : Vectorisation : les expressions de cas qui renvoient une valeur BOOLEAN ne sont pas prises en charge pour FILTER
Rétroportage	HIVE-20267 : Extension de WebUI pour inclure un formulaire afin de configurer dynamiquement les niveaux de journalisation

Type	Description
Rétroportage	HIVE-20263 : Faute de frappe dans la variable HiveReduceExpressionsWithStatsRule
Rétroportage	HIVE-20260 : Le NDV d'une colonne ne doit pas être redimensionné lorsque le nombre de lignes est modifié par un filtre sur une autre colonne
Rétroportage	HIVE-20252 : Réduction des semi-jointures : les cycles dus à une branche semi-jointive peuvent passer inaperçus si une jointure cartographique se trouve en amont d'une petite table.
Rétroportage	HIVE-20245 : Vectorisation : correction des problèmes de résultats nuls ou erronés dans BETWEEN/IN
Rétroportage	HIVE-20241 : Support des spécifications de partitionnement dans les instructions CTAS
Rétroportage	HIVE-20240 : Réduction des semi-jointures : utilisez une variable locale pour vérifier l'état de la table externe
Rétroportage	HIVE-20226 : HMS getNextNotification lancera une exception lorsque la requête MaxEvents dépasse le max_rows de la table
Rétroportage	HIVE-20225 : SerDe pour prendre en charge le format binaire Teradata
Rétroportage	HIVE-20213 : Mise à niveau de Calcite vers la version 1.17.0

Type	Description
Rétroportage	HIVE-20212 : HiveServer2 en mode http émet la métrique default.General.open_connections de manière incorrecte
Rétroportage	HIVE-20210 : L'optimiseur Simple Fetch devrait permettre de filtrer sur une colonne autre que la MapReduce partition et que la conversion sont minimales
Rétroportage	HIVE-20209 : La connexion au métastore échoue lors de la première tentative de vidage de repl
Rétroportage	HIVE-20207 : Vectorisation : correction des problèmes de résultats nuls ou erronés dans Filter/Comparer
Rétroportage	HIVE-20204 : Conversion de type pendant IN
Rétroportage	HIVE-20203 : Arrow divulgue un SerDe DirectByteBuffer
Rétroportage	HIVE-20197 : Vectorisation : ajout du test DECIMAL_64, ajout d'arithmétique et ajout d'autres tests GROUP BY Aggrégation Date/ Interval/Timestamp
Rétroportage	HIVE-20193 : CBOInfo n'est pas présent dans le fichier json du plan d'explication
Rétroportage	HIVE-20192 : HS2 avec le métastore intégré, des objets du gestionnaire fuient JDOPersis tence

Type	Description
Rétroportage	HIVE-20183 : L'insertion depuis une table à compartiments peut entraîner une perte de données si la table source contient un compartiment vide
Rétroportage	HIVE-20177 : Vectorisation : réduction de l'allocation en mode streaming KeyWrapper GroupBy
Rétroportage	HIVE-20174 : Vectorisation : correction des problèmes de résultats nuls ou erronés dans les fonctions d'agrégation GROUP BY
Rétroportage	HIVE-20172 : StatsUpdater échec avec une exception GSS lors de la tentative de connexion à une métastore distante
Rétroportage	HIVE-20153 : Count and Sum UDF consomment plus de mémoire dans Hive 2+
Rétroportage	HIVE-20152 : réinitialise l'état de la base de données, lorsque le vidage de repl échoue, afin de pouvoir renommer la table
Rétroportage	HIVE-20149 : échec/expiration du délai TestHiveCli
Rétroportage	HIVE-20130 : Meilleure journalisation pour le synchroniseur de schémas d'informations
Rétroportage	HIVE-20129 : Revenir à l'évolution du schéma basée sur la position pour les tables orc
Rétroportage	RUCHE 2018 :. SessionStateUserAuthenticator getGroupNames

Type	Description
Rétroportage	HIVE-20116 : utilise l'enregistreur TezTask parent
Rétroportage	HIVE-20115 : Les tableaux acides ne doivent pas utiliser le scan du pied de page pour l'analyse
Rétroportage	HIVE-20103 : WM : agréger les compteurs DAG uniquement si au moins un est utilisé
Rétroportage	HIVE-20101 : Bloom KFilter : évitez complètem ent d'utiliser les tableaux d'octets locaux []
Rétroportage	HIVE-20100 OpTraits : Select Optraits devrait cesser lorsqu'une incompatibilité est détectée
Rétroportage	HIVE-20098 : Statistiques : NPE lors de l'obtention des statistiques de partition de la colonne Date
Rétroportage	HIVE-20095 : Correction d'une fonctionnalité permettant de transférer le calcul vers des tables externes JDBC
Rétroportage	HIVE-20093 : LlapOutputFomatService : Utilisation ArrowBuf avec Netty pour la comptabilité
Rétroportage	HIVE-20090 : Étendre la création de filtres de réduction à semi-jointure afin de découvrir de nouvelles opportunités
Rétroportage	HIVE-20088 : Le chemin de localisation de la configuration Beeline est mal assemblé
Rétroportage	HIVE-20082 : la conversion HiveDecimal en chaîne ne formate pas correctement le décimal

Type	Description
Rétroportage	HIVE-20069 : Correction de la réoptimisation en cas d'optimisation DPP et Semijoin
Rétroportage	HIVE-20051 : Ignorer l'autorisation pour les tables temporaires
Rétroportage	HIVE-20044 : Arrow Serde doit enregistrer les valeurs des caractères et gérer correctement les chaînes vides
Rétroportage	HIVE-20028 : La configuration du cache du client Metastore n'est pas utilisée correctement
Rétroportage	HIVE-20025 : Nettoyage des fichiers d'événements créés par HiveProtoLoggingHook
Rétroportage	HIVE-20020 : Le fichier Jar Hive Contrib ne doit pas être dans lib
Rétroportage	HIVE-20013 : Ajouter un type de conversion implicite à date pour la fonction to_date
Rétroportage	HIVE-20011 : Élimination du mode ajout dans Proto Logging Hook
Rétroportage	HIVE-20005 : acid_table_stats, acid_no_buckets, etc. - modification du résultat de la requête sur la branche
Rétroportage	HIVE-20004 : Une mauvaise échelle utilisée par ConvertDecimal 64 ToDecimal entraîne des résultats incorrects
Rétroportage	HIVE-19995 : Trafic de lignes agrégé pour les tables acides

Type	Description
Rétroportage	HIVE-19993 : Il n'est pas possible d'utiliser un alias de table qui apparaît également sous forme de nom de colonne
Rétroportage	HIVE-19992 : Vectorisation : suivi de HIVE-19951 --> ajout d'un appel à. SchemaEvolution isOnlyImplicitConversion pour désactiver les E/S LLAP codées pour ORC uniquement lorsque la conversion de type de données n'est pas implicite
Rétroportage	HIVE-19989 : Metastore utilise un nom d'application incorrect pour les métriques HADOOP2
Rétroportage	HIVE-19981 : Les tables gérées converties en tables externes par l' HiveStrictManagedMigration utilitaire doivent être configurées pour supprimer les données lorsque la table est supprimée
Rétroportage	HIVE-19967 : Adhésion d'une PME : besoin d'Optraits pour une opération GBY PTFOperator
Rétroportage	HIVE-19935 : Session Hive WM interrompue : échec de la mise à jour du nombre de tâches LLAP
Rétroportage	HIVE-19924 : Marquer les tâches distcp exécutées par Repl Load
Rétroportage	HIVE-19891 : l'insertion dans des tables externes avec des répertoires de partitions personnalisés peut entraîner une perte de données

Type	Description
Rétroportage	HIVE-19850 : L'élimination des partitions dynamiques dans Tez entraîne l'erreur « Aucun travail trouvé pour TableScan »
Rétroportage	HIVE-19806 : Trier les résultats des tests pour éviter les défauts dans les résultats des tests
Rétroportage	HIVE-19770 : Support du CBO pour les requêtes comportant plusieurs colonnes identiques dans Select
Rétroportage	HIVE-19769 : Création d'objets dédiés pour les noms de bases de données et de tables
Rétroportage	HIVE-19765 : Ajouter des tests spécifiques au parquet à BlobstoreCliDriver
Rétroportage	HIVE-19759 : Test floconneux : # TestRpc testServerPort
Rétroportage	HIVE-19711 : Outil de schéma Refactor Hive
Rétroportage	HIVE-19701 : getDelegationToken FromMetaS tore n'a pas besoin d'être synchronisé
Rétroportage	HIVE-19694 : L'instruction Create Materialized View doit vérifier les conflits de noms de MV avant d'exécuter l'instruction SQL de MV.
Rétroportage	HIVE-19674 : Regrouper par constantes décimales jusqu'à la table Druid
Rétroportage	HIVE-19668 : Plus de 30 % du tas est gaspillé en dupliquant org.antlr.runtime. CommonTok enet chaînes dupliquées

Type	Description
Rétroportage	HIVE-19663 : génération de rapports LLAP IO refactorisés
Rétroportage	HIVE-19661 : changez Hive UDFs pour utiliser le moteur Re2J regex
Rétroportage	HIVE-19628 : NPE possible dans LLAP TestSigning
Rétroportage	HIVE-19568 : HS2 HA actif/passif : interdire la connexion directe à une instance passive HS2
Rétroportage	HIVE-19564 : Vectorisation : correction des problèmes de résultats nuls ou erronés en arithmétique
Rétroportage	HIVE-19552 : Activez #druidkafkamini_basic .q TestMiniDruidKafkaCliDriver
Rétroportage	HIVE-19432 : GetTablesOperation est trop lent si la ruche contient trop de bases de données et de tables
Rétroportage	HIVE-19360 : CBO : ajoute un « SQL optimisé » à l'objet QueryPlan
Rétroportage	HIVE-19326 : collecte automatique des statistiq ues : agrégation incorrecte lors des requêtes UNION
Rétroportage	HIVE-19313 : TestJdbcWith DBToken StoreNoDoAs les tests échouent
Rétroportage	HIVE-19285 : Ajoutez des journaux aux sous-classes de MetaDataOperation

Type	Description
Rétroportage	HIVE-19235 : Mettre à jour les fichiers dorés pour les tests Minimr
Rétroportage	HIVE-19104 : Lorsque le test MetaStore est lancé avec une nouvelle tentative, les instances doivent être indépendantes
Rétroportage	HIVE-18986 : Le changement de nom de table exécutera java.lang. StackOverflowError dans DataNucleus si la table contient un grand nombre de colonnes
Rétroportage	HIVE-18920 : CBO : initialise les fournisseurs Janino avant la première requête
Rétroportage	HIVE-18873 : Le fait d'ignorer silencieu sement le renvoi du prédicat pour la résonance magnétique HiveInputFormat peut entraîner la production de résultats erronés par les gestionnaires de stockage
Rétroportage	HIVE-18871 : Erreur d'exécution de hive on tez due à la définition de hive.aux.jars.path sur hdfs ://
Rétroportage	HIVE-18725 : Améliore la gestion des erreurs pour les sous-requêtes en cas de référence de colonne incorrecte
Rétroportage	HIVE-18696 : Les dossiers de partition risquent de ne pas être nettoyés correctement dans la méthode HiveMetaStore .add_partitions_core si un
Rétroportage	HIVE-18453 : ACID : ajout de la syntaxe « CREATE TRANSACTIONAL TABLE » pour unifier le support ACID ORC et Parquet

Type	Description
Rétroportage	HIVE-18201 : Désactiver XPROD_EDGE pour sq_count_chec
Rétroportage	HIVE-18140 : Les statistiques des tables partitionnées peuvent mal tourner dans les statistiques de base
Rétroportage	HIVE-17921 : L'agrégation avec une structure dans LLAP produit un résultat erroné
Rétroportage	HIVE-17896 : Top NKey : Création d'un opérateur Top vectorisable autonome NKey
Rétroportage	HIVE-17840 : HiveMetaStore mange une exception en cas d'échec de TransactionAlliste ners.NotifyEvent
Rétroportage	HIVE-17043 : Supprimer les colonnes non uniques du groupe par clés si elles ne sont pas référencées ultérieurement
Rétroportage	HIVE-17040 : Élimination conjointe en présence d'une relation FK
Rétroportage	HIVE-16839 : Appels déséquilibrés à OpenTransaction/CommitTransaction lors de la modification simultanée de la même partition
Rétroportage	HIVE-16100 : L'optimiseur de partitions triées dynamiques perd ses opérateurs frères
Rétroportage	HIVE-15956 : StackOverflowError lorsque vous supprimez de nombreuses partitions

Type	Description
Rétroportage	HIVE-15177 : L'authentification avec Hive échoue lorsque le type d'authentification Kerberos est défini sur FromSubject et que le principal contient _HOST
Rétroportage	HIVE-14898 : HS2 ne devrait pas enregistrer la pile d'appels pour une erreur d'en-tête d'authent ification vide
Rétroportage	HIVE-14493 : Support de partitionnement pour les vues matérialisées
Rétroportage	HIVE-14431 : Reconnaître COALESCE comme CASE
Rétroportage	HIVE-13457 : Création de points de terminaison d' HS2 API REST pour surveiller les informati ons
Rétroportage	HIVE-12342 : Définir la valeur par défaut de hive.optimize.index.filter sur true
Rétroportage	HIVE-10296 : Exception Cast observée lorsque Hive exécute une requête de jointure multiple sur Metastore
Rétroportage	HIVE-6980 : Supprimer une table en utilisant Direct SQL

Amazon EMR 6.6.0 - Modifications apportées à la configuration de Hive

- Dans le cadre de la modification OSS HIVE-20703, la propriété permettant de trier les partitions dynamiques, `hive.optimize.sort.dynamic.partition`, a été remplacée `hive.optimize.sort.dynamic.partition.threshold`.

  La configuration `hive.optimize.sort.dynamic.partition.threshold` possède les valeurs potentielles suivantes :

Valeur	Description
`0` (default)	Fait de l'optimisation du tri des partitions dynamiques une décision basée sur les coûts lorsque des fichiers ORC sont utilisés. Le nombre maximum d'écrivains autorisés dans les requêtes INSERT est calculé sur la base de (mémoire de l'exécuteur/du conteneur) * (pourcentage de mémoire utilisée par orc) divisé par la mémoire maximale (taille de bande) utilisée par un seul écrivain.
`-1`	Désactive l'optimisation pour trier complètement les partitions dynamiques.
`1`	Permet le tri global des partitions dynamiques. Cela permet de maintenir un seul enregistreur ouvert pour chaque valeur de partition dans le réducteur, réduisant ainsi la sollicitation de la mémoire sur les réducteurs.
`2` (ou nombre entier supérieur)	Indique à Hive d'utiliser l'entier spécifié comme seuil pour le nombre maximum de rédacteurs.

Amazon EMR 6.6.0 - Problèmes connus liés à Hive

- Les requêtes avec des fonctions de fenêtrage sur la même colonne que la jointure peuvent conduire à des transformations invalides comme rapporté dans HIVE-25278 et causer des résultats incorrects ou des échecs de requête. Pour contourner ce problème, vous pouvez désactiver CBO au niveau de la requête pour ce type de requête. Contactez AWS le support pour plus d'informations.
- Amazon EMR 6.6.0 inclut la version 3.1.2 du logiciel Hive. Hive 3.1.2 introduit une fonctionnalité qui divise les fichiers texte s'ils contiennent un en-tête et un pied de page (HIVE-21924). L'Apache Tez App Master lit chacun de vos fichiers pour déterminer les points de décalage dans la plage de données. Ces comportements combinés peuvent avoir un impact négatif sur les performances si vos requêtes lisent un grand nombre de petits fichiers texte. Pour contourner le problème, utilisez `CombineHiveInputFormat` et ajustez la taille de division maximale en configurant les propriétés suivantes :

```
SET hive.tez.input.format=org.apache.hadoop.hive.ql.io.CombineHiveInputFormat;
```

```
SET mapreduce.input.fileinputformat.split.maxsize=16777216;
```

- Avec Amazon EMR 6.6.0 à 6.9.x, les requêtes INSERT avec partition dynamique et clause ORDER BY ou SORT BY auront toujours deux réducteurs. Ce problème est dû à la modification d'OSS HIVE-20703, qui place l'optimisation des partitions dynamiques de tri dans le cadre d'une décision basée sur les coûts. Si votre charge de travail ne nécessite pas le tri des partitions dynamiques, nous vous recommandons de définir la propriété `hive.optimize.sort.dynamic.partition.threshold` sur `-1` pour désactiver la nouvelle fonctionnalité et obtenir le nombre de réducteurs correctement calculé. Ce problème est résolu dans OSS Hive dans le cadre de HIVE-22269 et dans Amazon EMR 6.10.0.

Hudi

Apache Hudi est une infrastructure de gestion des données open source utilisée pour simplifier le traitement des données incrémentielles et le développement de pipelines de données en fournissant des fonctionnalités d'insertion, de mise à jour, de mise à jour/insertion et de suppression au niveau de l'enregistrement. Mise à jour/insertion fait référence à la capacité à insérer des enregistrements dans un ensemble de données existant s'ils n'existent pas déjà ou de les mettre à jour s'ils existent. En gérant efficacement la façon dont les données sont présentées dans Amazon S3, Hudi permet d'ingérer et de mettre à jour des données quasiment en temps réel. Hudi gère soigneusement les métadonnées des actions effectuées sur l'ensemble de données pour faire en sorte que les actions soient atomiques et cohérentes.

Hudi est intégré à Apache Spark, Apache Hive et Presto. Dans les versions 6.1.0 et ultérieures d'Amazon EMR, Hudi est également intégré à Trino (PrestoSQL).

Avec la version 5.28.0 d'Amazon EMR et les versions ultérieures, EMR installe les composants Hudi par défaut lorsque Spark, Hive, Presto ou Flink sont installés. Vous pouvez utiliser Spark ou l' DeltaStreamer utilitaire Hudi pour créer ou mettre à jour des ensembles de données Hudi. Vous pouvez utiliser Hive, Spark, Presto ou Flink pour interroger un jeu de données Hudi de manière interactive ou créer des pipelines de traitement de données à l'aide de l'extraction incrémentielle. L'extraction incrémentielle fait référence à la possibilité d'extraire uniquement les données qui ont changé entre deux actions.

Ces fonctionnalités rendent Hudi adapté pour les cas d'utilisation suivants :

- Utilisation des données de streaming provenant de capteurs et d'autres appareils IoT (Internet des objets) nécessitant des événements d'insertion et de mise à jour spécifiques.
- Conformité aux règles de confidentialité des données dans des applications où les utilisateurs peuvent choisir d'être oubliés ou modifier leur consentement quant à la manière dont leurs données peuvent être utilisées.
- Implémentation d'un système de capture des données modifiées (CDC) qui vous permet d'appliquer des modifications à un ensemble de données au fil du temps.

Le tableau suivant répertorie la version de Hudi incluse dans la dernière version d’Amazon EMR série 7.x, ainsi que les composants qu’Amazon EMR installe avec Hudi.

Pour connaître la version des composants installés avec Hudi dans cette version, consultez Versions des composants de la version 7.9.0.

Informations sur la version de Hudi pour emr-7.9.0

Étiquette de version Amazon EMR	Version Hudi	Composants installés avec Hudi
emr-7.9.0	Hudi 0.15.0-amzn-6	Not available.

Le tableau suivant répertorie la version de Hudi incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Hudi.

Pour connaître la version des composants installés avec Hudi dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version de Hudi pour emr-6.15.0

Étiquette de version Amazon EMR	Version Hudi	Composants installés avec Hudi
emr-6.15.0	Hudi 0.14.0-amzn-0	Not available.

> **Note**
>
> La version 6.8.0 d'Amazon EMR est fournie avec Apache Hudi 0.11.1 ; toutefois, les clusters Amazon EMR 6.8.0 sont également compatibles avec le `hudi-spark3.3-bundle_2.12` open source de Hudi 0.12.0.

Le tableau suivant répertorie la version de Hudi incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Hudi.

Pour connaître la version des composants installés avec Hudi dans cette version, consultez Versions des composants de la version 5.36.2.

Informations sur la version de Hudi pour emr-5.36.2

Étiquette de version Amazon EMR	Version Hudi	Composants installés avec Hudi
emr-5.36.2	Hudi 0.10.1-amzn-1	Not available.

Rubriques

- Comment fonctionne Hudi
- Considérations et limitations pour l'utilisation de Hudi sur Amazon EMR
- Créez un cluster avec Hudi installé
- Travailler avec un jeu de données Hudi
- Utiliser la CLI Hudi
- Historique des sorties de Hudi

Comment fonctionne Hudi

Lorsque vous utilisez Hudi avec Amazon EMR, vous pouvez écrire des données dans le jeu de données à l'aide de l'API Spark Data Source ou de l' DeltaStreamer utilitaire Hudi. Hudi organise un jeu de données dans une structure de répertoire partitionnée sous *`basepath`* qui est similaire à une table Hive traditionnelle. Les spécificités de la façon dont les données sont présentées en tant que fichiers dans ces répertoires dépendent du type d'ensemble de données que vous choisissez. Vous pouvez choisir CoW (Copy on Write, copie sur écriture) ou MoR (Merge on Read, fusion sur lecture).

Quel que soit le type d'ensemble de données, chaque partition d'un ensemble de données est identifiée de manière unique par son chemin `partitionpath` relatif au chemin `basepath`. Au sein de chaque partition, les enregistrements sont répartis dans plusieurs fichiers de données. Pour de plus amples informations, veuillez consulter File management dans la documentation Apache Hudi.

Chaque action dans Hudi est associée à un commit correspondant, identifié par un horodatage croissant de façon monotone appelé Instant. Hudi conserve une série de toutes les actions effectuées sur le jeu de données sous forme de chronologie. Hudi s'appuie sur la chronologie pour isoler instantanément les lecteurs et les rédacteurs, et pour permettre de revenir à un point antérieur dans le temps. Pour plus d'informations sur les actions enregistrées par Hudi et l'état des actions, consultez Timeline dans la documentation Apache Hudi.

Comprendre les types de stockage des jeux de données : copie en écriture ou fusion en lecture

Lorsque vous créez un jeu de données Hudi, vous spécifiez que celui-ci doit être copié en écriture ou fusionné en lecture.

- Copie sur écriture (CoW) – Les données sont stockées dans un format en colonnes (Parquet) et chaque mise à jour crée une nouvelle version des fichiers lors d'une écriture. CoW est le type de stockage par défaut.
- Fusion sur lecture (MoR) – Les données sont stockées en utilisant une combinaison de formats en colonnes (Parquet) et en lignes (Avro). Les mises à jour sont consignées dans des fichiers delta basés sur les lignes et sont compressées si nécessaire pour créer de nouvelles versions des fichiers en colonnes.

Avec les ensembles de données CoW, chaque fois qu'une mise à jour est apportée à un enregistrement, le fichier qui contient l'enregistrement est réécrit avec les valeurs mises à jour. Avec un jeu de données MoR, chaque fois qu'une mise à jour a lieu, Hudi écrit uniquement la ligne du registre modifié. Le type de stockage MoR est mieux adapté aux charges de travail donnant lieu à de nombreuses écritures ou modifications avec moins de lectures. Le type de stockage CoW est mieux adapté aux charges de travail donnant lieu à de nombreuses lectures sur des données qui changent moins souvent.

Hudi fournit trois vues logiques pour l'accès aux données :

- Vue optimisée en lecture – Fournit le dernier ensemble de données validé de tables CoW et le dernier jeu de données compressé de tables MoR.
- Vue incrémentielle – Fournit un flux de modification entre deux actions d'un ensemble de données CoW pour alimenter les tâches et les flux de travail ETL (extraction, transformation et chargement) en aval.
- Vue en temps réel – Fournit les dernières données validées d'une table de MoR en fusionnant en ligne les fichiers de colonnes et les fichiers basés sur des lignes.

Lorsque vous interrogez la vue optimisée en lecture, la requête renvoie toutes les données compressées mais n'inclut pas les dernières validations delta. L'interrogation de ces données fournit de bonnes performances de lecture, mais ignore les données les plus à jour. Lorsque vous interrogez la vue en temps réel, Hudi fusionne les données compressées avec les validations delta sur lecture.

Les données les plus à jour sont disponibles pour interrogation, mais la surcharge de calcul de la fusion rend la requête moins performante. La possibilité d'interroger des données compressées ou des données en temps réel vous permet de choisir entre performances et flexibilité lorsque vous exécutez des requêtes.

Pour plus d'informations sur les avantages et inconvénients des types de stockage, consultez Storage Types & Views dans la documentation Apache Hudi.

Hudi crée deux tables dans le metastore Hive pour le type de stockage MoR : une table avec le nom que vous avez spécifié, qui est une vue optimisée en lecture, et une table avec le même nom auquel est ajouté `_rt`, qui est une vue en temps réel. Vous pouvez interroger les deux tables.

Enregistrement d'un jeu de données Hudi auprès de votre métastore

Lorsque vous enregistrez une table Hudi auprès du metastore Hive, vous pouvez interroger des tables Hudi à l'aide de Hive, Spark SQL ou Presto comme vous le feriez pour n'importe quelle autre table. En outre, vous pouvez intégrer Hudi à AWS Glue en configurant Hive et Spark pour qu'ils utilisent le catalogue de données AWS Glue comme métastore. Pour les tables MoR, Hudi enregistre le jeu de données sous la forme de deux tables dans le metastore : une table avec le nom que vous avez spécifié, qui est une vue optimisée en lecture, et une table avec le même nom auquel est ajouté `_rt`, qui est une vue en temps réel.

Vous enregistrez une table Hudi auprès du metastore Hive lorsque vous utilisez Spark pour créer un jeu de données Hudi en définissant l'option `HIVE_SYNC_ENABLED_OPT_KEY` sur `"true"` et en fournissant d'autres propriétés requises. Pour de plus amples informations, veuillez consulter Travailler avec un jeu de données Hudi. En outre, vous pouvez utiliser l'utilitaire de ligne de commande hive_sync_tool pour enregistrer un jeu de données Hudi en tant que table dans votre metastore, séparément.

Considérations et limitations pour l'utilisation de Hudi sur Amazon EMR

- Le champ de clé d'enregistrement ne peut pas être nul ou vide – Le champ que vous spécifiez comme champ de clé d'enregistrement ne peut pas avoir des valeurs `null` ou vides.
- Schéma mis à jour par défaut lors de l'upsert et de l'insert — Hudi fournit une interface `HoodieRecordPayload` qui détermine comment le jeu de données Hudi d'entrée DataFrame et existant sont fusionnés pour produire un nouveau jeu de données mis à jour. Hudi fournit une implémentation par défaut de cette classe`OverwriteWithLatestAvroPayload`, qui remplace

les enregistrements existants et met à jour le schéma tel que spécifié dans l'entrée. DataFrame Pour personnaliser cette logique en vue d'une implémentation de la fusion et des mises à jour partielles, vous pouvez fournir une implémentation de l'interface `HoodieRecordPayload` à l'aide du paramètre `DataSourceWriteOptions.PAYLOAD_CLASS_OPT_KEY`.

- La suppression exige un schéma – Lors d'une suppression, vous devez spécifier les champs de clé d'enregistrement, de clé de partition et de clé de pré-combinaison. Les autres colonnes peuvent être `null` ou vides, mais le schéma complet est requis.
- Limites des tables MoR – Les tables MoR ne prennent pas en charge le pointage de sauvegarde. Vous pouvez interroger les tables MoR à l'aide de la vue optimisée pour la lecture ou de la vue en temps réel (*`tableName`*`_rt`) de Spark SQL, Presto ou Hive. L'utilisation de la vue optimisée pour la lecture expose uniquement les données du fichier de base et n'expose pas une vue fusionnée des données de base et des données de journal.
- Hive
  - Pour l'enregistrement des tables dans le metastore Hive, Hudi s'attend à ce que le serveur Hive Thrift s'exécute sur le port par défaut `10000`. Si vous remplacez ce port par un port personnalisé, transmettez l'option `HIVE_URL_OPT_KEY` comme indiqué dans l'exemple suivant.

    ```
    .option(DataSourceWriteOptions.HIVE_URL_OPT_KEY, "jdbc:hive2://localhost:override-port-number
    ```

  - Le type de données `timestamp` dans Spark est enregistré comme type de données `long` dans Hive, et non comme type `timestamp` de Hive.
- Presto
  - Presto ne prend pas en charge la lecture des tables MoR en temps réel dans les versions de Hudi inférieures à 0.6.0.
  - Presto ne prend en charge que les requêtes instantanées.
  - Pour que Presto interprète correctement les colonnes d'un ensemble de données, Hudi définissez la valeur de `hive.parquet_use_column_names` sur `true`.
    - Pour définir la valeur pour une session, dans le shell Presto, exécutez la commande suivante :

      ```
      set session hive.parquet_use_column_names=true
      ```

    - Pour définir la valeur au niveau du cluster, utilisez la classification de configuration `presto-connector-hive` pour définir `hive.parquet.use_column_names` sur `true`, comme illustré dans l'exemple suivant. Pour de plus amples informations, veuillez consulter Configuration des applications.

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.parquet.use-column-names": "true"
    }
  }
]
```

- HBase Index
  - La HBase version utilisée pour créer Hudi peut être différente de celle répertoriée dans le guide de publication de l'EMR. Pour extraire les dépendances correctes pour votre session Spark, exécutez la commande suivante.

```
spark-shell \
--jars /usr/lib/spark/external/lib/spark-avro.jar,/usr/lib/hudi/cli/lib/*.jar \
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \
--conf "spark.sql.hive.convertMetastoreParquet=false"
```

- Paramètres pour de meilleures performances — Pour EMR 7.3+/Hudi 0.15+, il est recommandé aux clients de définir cette configuration afin de réduire la charge de sérialisation de Kryo :

```
--conf 'spark.kryo.registrator=org.apache.spark.HoodieKryoRegistrar'
```

> **Note**
>
> Si vous utilisez un contrôle d'accès fin (FGAC) sur EMR Serverless, cette configuration n'est pas nécessaire, car les utilisateurs doivent utiliser plutôt que. JavaSerializer KryoSerializer

Créez un cluster avec Hudi installé

Avec Amazon EMR version 5.28.0 ou ultérieure, Amazon EMR installe les composants Hudi par défaut lorsque Spark, Hive ou Presto est installé. Pour utiliser Hudi sur Amazon EMR, créez un cluster avec une ou plusieurs des applications suivantes installées :

- Hadoop

- Hive
- Spark
- Presto
- Flink

Vous pouvez créer un cluster à l'aide da la AWS Management Console AWS CLI, de l'ou de l'API Amazon EMR.

Création d'un cluster avec Hudi à l'aide d' AWS Management Console

1. Accédez à la nouvelle console Amazon EMR et sélectionnez Changer pour l'ancienne console depuis le menu latéral. Pour plus d'informations sur ce qu'implique le passage à l'ancienne console, consultez la rubrique Utilisation de l'ancienne console.
2. Choisissez Créer un cluster et Go to advanced options (Aller aux options avancées).
3. Sous Software Configuration (Configuration logicielle), choisissez emr-5.28.0 ou version ultérieure pour Release (Version) et sélectionnez Hadoop, Hive, Spark, Presto et Tez avec les autres applications dont votre cluster a besoin.
4. Configurez d'autres options si nécessaire pour votre application, puis choisissez Next (Suivant).
5. Configurez les options pour Hardware (Matériel) and General cluster settings (Paramètres de cluster généraux) à votre guise.
6. Pour les options de sécurité, nous vous recommandons de sélectionner une paire de EC2 clés que vous pouvez utiliser pour vous connecter à la ligne de commande du nœud principal via SSH. Cela vous permet d'exécuter les commandes du shell Spark, les commandes de l'interface de ligne de commande Hive et les commandes de l'interface de ligne de commande Hudi décrites dans ce guide.
7. Configurez les autres options de sécurité comme vous le souhaitez et choisissez Create cluster (Créer un cluster).

Travailler avec un jeu de données Hudi

Hudi prend en charge l'insertion, la mise à jour et la suppression de données dans des jeux de données Hudi via Spark. Pour plus d'informations, consultez Writing Hudi tables dans la documentation Apache Hudi.

Les exemples suivants démontrent comment lancer le shell Spark interactif, utiliser Spark submit ou utiliser Blocs-notes EMR pour travailler avec Hudi sur Amazon EMR. Vous pouvez également utiliser l' DeltaStreamer utilitaire Hudi ou d'autres outils pour écrire dans un ensemble de données. Dans cette section, les exemples illustrent l'utilisation des ensembles de données à l'aide du shell Spark lorsque vous êtes connecté au nœud principal à l'aide de SSH comme utilisateur `hadoop` par défaut.

Lancez le shell Spark à l'aide d'Amazon EMR 6.7 et versions ultérieures

Lorsque vous exécutez `spark-shell`, `spark-submit` ou `spark-sql` ou utilisez Amazon EMR 6.7.0 ou version ultérieure, passez les commandes suivantes.

> **Note**
>
> Amazon EMR 6.7.0 utilise Apache Hudi 0.11.0-amzn-0, qui contient des améliorations significatives par rapport aux versions précédentes de Hudi. Pour plus d'informations, consultez le Guide de migration Apache Hudi 0.11.0. Les exemples de cet onglet reflètent ces changements.

Pour ouvrir le shell Spark sur le nœud primaire

1. Connexion au nœud primaire à l'aide de SSH. Pour plus d'informations, consultez Connexion au nœud primaire à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
2. Entrez la commande suivante pour lancer le shell Spark. Pour utiliser la PySpark coque, remplacez-la *`spark-shell`* par*`pyspark`*.

```
spark-shell --jars /usr/lib/hudi/hudi-spark-bundle.jar \
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \
--conf
 "spark.sql.catalog.spark_catalog=org.apache.spark.sql.hudi.catalog.HoodieCatalog"
 \
--conf "spark.sql.extensions=org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
```

Lancez le shell Spark à l'aide d'Amazon EMR 6.6 et versions antérieures

Lors de l'exécution `spark-shell`, `spark-submit` ou `spark-sql` ou utilisez d'Amazon EMR 6.6.x ou version antérieure, passez les commandes suivantes.

> **Note**
>
> - Amazon EMR 6.2 et 5.31 et versions ultérieures (Hudi 0.6.x et versions ultérieures) peuvent omettre `spark-avro.jar` de la configuration.
> - Amazon EMR 6.5 et 5.35 et versions ultérieures (Hudi 0.9.x et versions ultérieures) peuvent omettre `spark.sql.hive.convertMetastoreParquet=false` de la configuration.
> - Amazon EMR 6.6 et 5.36 et versions ultérieures (Hudi 0.10.x et versions ultérieures) doivent inclure la configuration `HoodieSparkSessionExtension` telle que décrite dans le Guide Spark de la version : 0.10.0 :
>
> ```
> --conf
>  "spark.sql.extensions=org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
>  \
> ```

Pour ouvrir le shell Spark sur le nœud primaire

1. Connexion au nœud primaire à l'aide de SSH. Pour plus d'informations, consultez Connexion au nœud primaire à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
2. Entrez la commande suivante pour lancer le shell Spark. Pour utiliser la PySpark coque, remplacez-la *`spark-shell`* par*`pyspark`*.

```
spark-shell \
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \
--conf "spark.sql.hive.convertMetastoreParquet=false" \
--jars /usr/lib/hudi/hudi-spark-bundle.jar,/usr/lib/spark/external/lib/spark-
avro.jar
```

Utiliser Hudi avec Blocs-notes Amazon EMR utilisant Amazon EMR 6.7 et versions ultérieures

Pour utiliser Hudi avec les blocs-notes Amazon EMR, vous devez d'abord copier les fichiers jar Hudi depuis le système de fichiers local vers HDFS sur le nœud principal du cluster de bloc-notes. Vous pouvez utiliser l'éditeur de bloc-notes pour configurer votre bloc-notes EMR afin d'utiliser Hudi.

Pour utiliser Hudi avec Blocs-notes Amazon EMR

1. Créez et lancez un cluster pour le Blocs-notes Amazon EMR. Pour plus d'informations, consultez Création de clusters Amazon EMR pour les blocs-notes dans le Guide de gestion Amazon EMR.
2. Connectez-vous au nœud principal du cluster à l'aide de SSH, puis copiez les fichiers jar depuis le système de fichiers local vers HDFS comme illustré dans les exemples suivants. Dans l'exemple, nous créons un répertoire dans HDFS pour plus de clarté dans la gestion des fichiers. Vous pouvez choisir votre propre destination dans HDFS si vous le souhaitez.

```
hdfs dfs -mkdir -p /apps/hudi/lib
```

```
hdfs dfs -copyFromLocal /usr/lib/hudi/hudi-spark-bundle.jar /apps/hudi/lib/hudi-
spark-bundle.jar
```

3. Ouvrez l'éditeur de bloc-notes, entrez le code de l'exemple suivant et exécutez-le.

```
%%configure
{ "conf": {
            "spark.jars":"hdfs:///apps/hudi/lib/hudi-spark-bundle.jar",
            "spark.serializer":"org.apache.spark.serializer.KryoSerializer",
            "spark.sql.catalog.spark_catalog":
 "org.apache.spark.sql.hudi.catalog.HoodieCatalog",

 "spark.sql.extensions":"org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
          }}
```

Utiliser Hudi avec Blocs-notes Amazon EMR utilisant Amazon EMR 6.6 et versions antérieures

Pour utiliser Hudi avec les blocs-notes Amazon EMR, vous devez d'abord copier les fichiers jar Hudi depuis le système de fichiers local vers HDFS sur le nœud principal du cluster de bloc-notes. Vous pouvez utiliser l'éditeur de bloc-notes pour configurer votre bloc-notes EMR afin d'utiliser Hudi.

Pour utiliser Hudi avec Blocs-notes Amazon EMR

1. Créez et lancez un cluster pour le Blocs-notes Amazon EMR. Pour plus d'informations, consultez Création de clusters Amazon EMR pour les blocs-notes dans le Guide de gestion Amazon EMR.

2. Connectez-vous au nœud principal du cluster à l'aide de SSH, puis copiez les fichiers jar depuis le système de fichiers local vers HDFS comme illustré dans les exemples suivants. Dans l'exemple, nous créons un répertoire dans HDFS pour plus de clarté dans la gestion des fichiers. Vous pouvez choisir votre propre destination dans HDFS si vous le souhaitez.

```
hdfs dfs -mkdir -p /apps/hudi/lib
```

```
hdfs dfs -copyFromLocal /usr/lib/hudi/hudi-spark-bundle.jar /apps/hudi/lib/hudi-
spark-bundle.jar
```

```
hdfs dfs -copyFromLocal /usr/lib/spark/external/lib/spark-avro.jar /apps/hudi/lib/
spark-avro.jar
```

3. Ouvrez l'éditeur de bloc-notes, entrez le code de l'exemple suivant et exécutez-le.

```
{ "conf": {
            "spark.jars":"hdfs:///apps/hudi/lib/hudi-spark-bundle.jar,hdfs:///apps/
hudi/lib/spark-avro.jar",
            "spark.serializer":"org.apache.spark.serializer.KryoSerializer",
            "spark.sql.hive.convertMetastoreParquet":"false"
          }}
```

Initialisation d'une session Spark pour Hudi

Lorsque vous utilisez Scala, vous devez importer les classes suivantes dans votre session Spark. Vous devez le faire une fois par session Spark.

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._
import org.apache.hudi.DataSourceWriteOptions
import org.apache.hudi.DataSourceReadOptions
import org.apache.hudi.config.HoodieWriteConfig
import org.apache.hudi.hive.MultiPartKeysValueExtractor
import org.apache.hudi.hive.HiveSyncConfig
import org.apache.hudi.sync.common.HoodieSyncConfig
```

Écrire dans un jeu de données Hudi

Les exemples suivants montrent comment créer un jeu de données Hudi DataFrame et l'écrire en tant que jeu de données Hudi.

> **Note**
>
> Pour coller des exemples de code dans le shell Spark, tapez **`:paste`** à l'invite, collez l'exemple, puis appuyez sur **`CTRL`** + **`D`**.

Chaque fois que vous écrivez un dans DataFrame un jeu de données Hudi, vous devez spécifier`DataSourceWriteOptions`. Beaucoup de ces options sont susceptibles d'être identiques dans les opérations d'écriture. L'exemple suivant spécifie les options communes à l'aide de la variable *`hudiOptions`*, que les exemples suivants utilisent.

Écrire à l'aide de Scala avec Amazon EMR 6.7 et versions ultérieures

> **Note**
>
> Amazon EMR 6.7.0 utilise Apache Hudi 0.11.0-amzn-0, qui contient des améliorations significatives par rapport aux versions précédentes de Hudi. Pour plus d'informations, consultez le Guide de migration Apache Hudi 0.11.0. Les exemples de cet onglet reflètent ces changements.

```
// Create a DataFrame
val inputDF = Seq(
 ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
 ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
 ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
 ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
 ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
 ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z")
 ).toDF("id", "creation_date", "last_update_time")

//Specify common DataSourceWriteOptions in the single hudiOptions variable
val hudiOptions = Map[String,String](
  HoodieWriteConfig.TBL_NAME.key -> "tableName",
  DataSourceWriteOptions.TABLE_TYPE.key -> "COPY_ON_WRITE",
```

```
  DataSourceWriteOptions.RECORDKEY_FIELD_OPT_KEY -> "id",
  DataSourceWriteOptions.PARTITIONPATH_FIELD_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.PRECOMBINE_FIELD_OPT_KEY -> "last_update_time",
  DataSourceWriteOptions.HIVE_SYNC_ENABLED_OPT_KEY -> "true",
  DataSourceWriteOptions.HIVE_TABLE_OPT_KEY -> "tableName",
  DataSourceWriteOptions.HIVE_PARTITION_FIELDS_OPT_KEY -> "creation_date",
  HoodieSyncConfig.META_SYNC_PARTITION_EXTRACTOR_CLASS.key ->
 "org.apache.hudi.hive.MultiPartKeysValueExtractor",
  HoodieSyncConfig.META_SYNC_ENABLED.key -> "true",
  HiveSyncConfig.HIVE_SYNC_MODE.key -> "hms",
  HoodieSyncConfig.META_SYNC_TABLE_NAME.key -> "tableName",
  HoodieSyncConfig.META_SYNC_PARTITION_FIELDS.key -> "creation_date"
)

// Write the DataFrame as a Hudi dataset
(inputDF.write
    .format("hudi")
    .options(hudiOptions)
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,"insert")
    .mode(SaveMode.Overwrite)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Écrire à l'aide de Scala avec Amazon EMR 6.6 et versions antérieures

```
// Create a DataFrame
val inputDF = Seq(
 ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
 ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
 ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
 ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
 ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
 ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z")
 ).toDF("id", "creation_date", "last_update_time")

//Specify common DataSourceWriteOptions in the single hudiOptions variable
val hudiOptions = Map[String,String](
  HoodieWriteConfig.TABLE_NAME -> "tableName",
  DataSourceWriteOptions.TABLE_TYPE_OPT_KEY -> "COPY_ON_WRITE",
  DataSourceWriteOptions.RECORDKEY_FIELD_OPT_KEY -> "id",
  DataSourceWriteOptions.PARTITIONPATH_FIELD_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.PRECOMBINE_FIELD_OPT_KEY -> "last_update_time",
  DataSourceWriteOptions.HIVE_SYNC_ENABLED_OPT_KEY -> "true",
  DataSourceWriteOptions.HIVE_TABLE_OPT_KEY -> "tableName",
```

```
  DataSourceWriteOptions.HIVE_PARTITION_FIELDS_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY ->
 classOf[MultiPartKeysValueExtractor].getName
)

// Write the DataFrame as a Hudi dataset
(inputDF.write
    .format("org.apache.hudi")
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
 DataSourceWriteOptions.INSERT_OPERATION_OPT_VAL)
    .options(hudiOptions)
    .mode(SaveMode.Overwrite)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Écrivez en utilisant PySpark

```
# Create a DataFrame
inputDF = spark.createDataFrame(
    [
        ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
        ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
        ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
        ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
        ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
        ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z"),
    ],
    ["id", "creation_date", "last_update_time"]
)

# Specify common DataSourceWriteOptions in the single hudiOptions variable
hudiOptions = {
'hoodie.table.name': 'tableName',
'hoodie.datasource.write.recordkey.field': 'id',
'hoodie.datasource.write.partitionpath.field': 'creation_date',
'hoodie.datasource.write.precombine.field': 'last_update_time',
'hoodie.datasource.hive_sync.enable': 'true',
'hoodie.datasource.hive_sync.table': 'tableName',
'hoodie.datasource.hive_sync.partition_fields': 'creation_date',
'hoodie.datasource.hive_sync.partition_extractor_class':
 'org.apache.hudi.hive.MultiPartKeysValueExtractor'
}

# Write a DataFrame as a Hudi dataset
```

```
inputDF.write \
.format('org.apache.hudi') \
.option('hoodie.datasource.write.operation', 'insert') \
.options(**hudiOptions) \
.mode('overwrite') \
.save('s3://amzn-s3-demo-bucket/myhudidataset/')
```

> **Note**
>
> Vous pouvez voir « hoodie » au lieu de Hudi dans les exemples de code et les notifications. La base de code Hudi utilise largement l'ancienne orthographe « hoodie ».

DataSourceWriteOptions référence pour Hudi

Option	Description
TABLE_NAME	Nom de la table sous laquelle enregistrer l'ensemble de données.
TABLE_TYPE_OPT_KEY	Facultatif. Spécifie si l'ensemble de données est créé en tant que `"COPY_ON_WRITE"` ou `"MERGE_ON_READ"` . L'argument par défaut est `"COPY_ON_WRITE"` .
RECORDKEY_FIELD_OPT_KEY	Champ de clé d'enregistrement dont la valeur sera utilisée en tant que composant `recordKey` de `HoodieKey` . La valeur réelle sera obtenue en appelant `.toString()` sur la valeur de champ. Des champs imbriqués peuvent être spécifiés à l'aide de la notation par points, par exemple, `a.b.c`.
PARTITIONPATH_FIELD_OPT_KEY	Champ de chemin de partition dont la valeur sera utilisée en tant que composant `partitionPath` de `HoodieKey` . La valeur réelle sera obtenue en appelant `.toString()` sur la valeur de champ.

Option	Description
PRECOMBINE_FIELD_OPT_KEY	Champ utilisé dans la pré-combinaison avant l'écriture réelle. Lorsque deux enregistrements ont la même valeur de clé, Hudi sélectionne celui avec la plus grande valeur pour le champ de pré-combinaison, comme déterminé par `Object.compareTo(..)` .

Les options suivantes sont nécessaires uniquement pour enregistrer la table du jeu de données Hudi dans votre metastore. Si vous n'enregistrez pas votre jeu de données Hudi en tant que table dans le metastore Hive, ces options ne sont pas requises.

DataSourceWriteOptions référence pour Hive

Option	Description
HIVE_DATABASE_OPT_KEY	Base de données Hive avec laquelle synchroniser. L'argument par défaut est `"default"` .
HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY	Classe utilisée pour extraire les valeurs de champ de partition dans les colonnes de partition Hive.
HIVE_PARTITION_FIELDS_OPT_KEY	Champ dans l'ensemble de données à utiliser pour déterminer les colonnes de partition Hive.
HIVE_SYNC_ENABLED_OPT_KEY	Lorsqu'elle est définie sur `"true"`, enregistre l'ensemble de données auprès du metastore Apache Hive. L'argument par défaut est `"false"`.
HIVE_TABLE_OPT_KEY	Obligatoire. Nom de la table dans Hive avec laquelle synchroniser. Par exemple, `"my_hudi_table_cow"` .
HIVE_USER_OPT_KEY	Facultatif. Nom d'utilisateur Hive à utiliser lors de la synchronisation. Par exemple, `"hadoop"`.

Option	Description
HIVE_PASS_OPT_KEY	Facultatif. Mot de passe Hive pour l'utilisateur spécifié par `HIVE_USER_OPT_KEY` .
HIVE_URL_OPT_KEY	URL du metastore Hive.

Données d'Upsert

L'exemple suivant montre comment insérer des données en écrivant un DataFrame. Contrairement à l'exemple d'insertion précédent, la valeur `OPERATION_OPT_KEY` est définie sur `UPSERT_OPERATION_OPT_VAL`. En outre, `.mode(SaveMode.Append)` est spécifié pour indiquer que l'enregistrement doit être ajouté.

Upsert à l'aide de Scala avec Amazon EMR 6.7 et versions ultérieures

> **Note**
>
> Amazon EMR 6.7.0 utilise Apache Hudi 0.11.0-amzn-0, qui contient des améliorations significatives par rapport aux versions précédentes de Hudi. Pour plus d'informations, consultez le Guide de migration Apache Hudi 0.11.0. Les exemples de cet onglet reflètent ces changements.

```
// Create a new DataFrame from the first row of inputDF with a different creation_date
 value
val updateDF = inputDF.limit(1).withColumn("creation_date", lit("new_value"))

(updateDF.write
    .format("hudi")
    .options(hudiOptions)
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY, "upsert")
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Upsert en utilisant Scala avec Amazon EMR 6.6 et versions antérieures

```
// Create a new DataFrame from the first row of inputDF with a different creation_date
 value
val updateDF = inputDF.limit(1).withColumn("creation_date", lit("new_value"))

(updateDF.write
    .format("org.apache.hudi")
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
 DataSourceWriteOptions.UPSERT_OPERATION_OPT_VAL)
    .options(hudiOptions)
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Upsert en utilisant PySpark

```
from pyspark.sql.functions import lit

# Create a new DataFrame from the first row of inputDF with a different creation_date
 value
updateDF = inputDF.limit(1).withColumn('creation_date', lit('new_value'))

updateDF.write \
    .format('org.apache.hudi') \
    .option('hoodie.datasource.write.operation', 'upsert') \
    .options(**hudiOptions) \
    .mode('append') \
    .save('s3://amzn-s3-demo-bucket/myhudidataset/')
```

Supprimer un enregistrement

Pour supprimer définitivement un enregistrement, vous pouvez insérer une charge utile vide. Dans ce cas, l'option `PAYLOAD_CLASS_OPT_KEY` spécifie la classe `EmptyHoodieRecordPayload`. L'exemple utilise le même DataFrame`updateDF`, que celui utilisé dans l'exemple d'upsert pour spécifier le même enregistrement.

Supprimer à l'aide de Scala avec Amazon EMR 6.7 et versions ultérieures

> **Note**
>
> Amazon EMR 6.7.0 utilise Apache Hudi 0.11.0-amzn-0, qui contient des améliorations significatives par rapport aux versions précédentes de Hudi. Pour plus d'informations,

consultez le Guide de migration Apache Hudi 0.11.0. Les exemples de cet onglet reflètent ces changements.

```
(updateDF.write
    .format("hudi")
    .options(hudiOptions)
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY, "delete")
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Supprimer à l'aide de Scala avec Amazon EMR 6.6 et versions antérieures

```
(updateDF.write
    .format("org.apache.hudi")
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
 DataSourceWriteOptions.UPSERT_OPERATION_OPT_VAL)
    .option(DataSourceWriteOptions.PAYLOAD_CLASS_OPT_KEY,
 "org.apache.hudi.common.model.EmptyHoodieRecordPayload")
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Supprimer en utilisant PySpark

```
updateDF.write \
    .format('org.apache.hudi') \
    .option('hoodie.datasource.write.operation', 'upsert') \
    .option('hoodie.datasource.write.payload.class',
 'org.apache.hudi.common.model.EmptyHoodieRecordPayload') \
    .options(**hudiOptions) \
    .mode('append') \
    .save('s3://amzn-s3-demo-bucket/myhudidataset/')
```

Vous pouvez également supprimer définitivement des données en définissant `OPERATION_OPT_KEY` sur `DELETE_OPERATION_OPT_VAL` pour supprimer tous les enregistrements de le jeu de données que vous soumettez. Pour obtenir des instructions sur les suppressions logicielles et pour plus d'informations sur la suppression de données stockées dans des tables Hudi, consultez Suppressions dans la documentation d'Apache Hudi.

Lire à partir d'un jeu de données Hudi

Pour récupérer les données à l'heure actuelle, Hudi effectue des requêtes instantanées par défaut. Voici un exemple d'interrogation de le jeu de données écrit dans S3 dans Écrire dans un jeu de données Hudi. Remplacez *s3://amzn-s3-demo-bucket/myhudidataset* par le chemin de votre table et ajoutez des astérisques génériques pour chaque niveau de partition, ainsi qu'un astérisque supplémentaire. Dans cet exemple, il existe un niveau de partition. Nous avons donc ajouté deux symboles génériques.

Lire à l'aide de Scala avec Amazon EMR 6.7 et versions ultérieures

> **Note**
>
> Amazon EMR 6.7.0 utilise Apache Hudi 0.11.0-amzn-0, qui contient des améliorations significatives par rapport aux versions précédentes de Hudi. Pour plus d'informations, consultez le Guide de migration Apache Hudi 0.11.0. Les exemples de cet onglet reflètent ces changements.

```
val snapshotQueryDF = spark.read
    .format("hudi")
    .load("s3://amzn-s3-demo-bucket/myhudidataset")
    .show()
```

Lire à l'aide de Scala avec Amazon EMR 6.6 et versions antérieures

```
(val snapshotQueryDF = spark.read
    .format("org.apache.hudi")
    .load("s3://amzn-s3-demo-bucket/myhudidataset" + "/*/*"))

snapshotQueryDF.show()
```

Lisez en utilisant PySpark

```
snapshotQueryDF = spark.read \
    .format('org.apache.hudi') \
    .load('s3://amzn-s3-demo-bucket/myhudidataset' + '/*/*')

snapshotQueryDF.show()
```

Requêtes incrémentielles

Vous pouvez également effectuer des requêtes incrémentielles avec Hudi pour obtenir un flux d'enregistrements modifiés depuis un horodatage de validation donné. Pour ce faire, définissez le champ `QUERY_TYPE_OPT_KEY` sur `QUERY_TYPE_INCREMENTAL_OPT_VAL`. Ajoutez ensuite une valeur pour `BEGIN_INSTANTTIME_OPT_KEY` pour obtenir tous les enregistrements écrits depuis l'heure spécifiée. Les requêtes incrémentielles sont généralement dix fois plus efficaces que leurs homologues par lots, car elles ne traitent que les enregistrements modifiés.

Lorsque vous effectuez des requêtes incrémentielles, utilisez le chemin de la table racine (de base) sans les astérisques génériques utilisés pour les requêtes Snapshot.

> **Note**
>
> Presto ne prend pas en charge les requêtes incrémentielles.

Requêtes incrémentielles à l'aide de Scala

```
(val incQueryDF = spark.read
    .format("org.apache.hudi")
    .option(DataSourceReadOptions.QUERY_TYPE_OPT_KEY,
 DataSourceReadOptions.QUERY_TYPE_INCREMENTAL_OPT_VAL)
    .option(DataSourceReadOptions.BEGIN_INSTANTTIME_OPT_KEY, <beginInstantTime>)
    .load("s3://amzn-s3-demo-bucket/myhudidataset" ))

incQueryDF.show()
```

Requêtes incrémentielles à l'aide d' PySpark

```
readOptions = {
  'hoodie.datasource.query.type': 'incremental',
  'hoodie.datasource.read.begin.instanttime': <beginInstantTime>,
}

incQueryDF = spark.read \
    .format('org.apache.hudi') \
    .options(**readOptions) \
    .load('s3://amzn-s3-demo-bucket/myhudidataset')
```

```
incQueryDF.show()
```

Pour plus d'informations sur la lecture d'jeux de données Hudi, consultez la rubrique Interrogation de tables Hudi dans la documentation d'Apache Hudi.

Utiliser la CLI Hudi

Vous pouvez utiliser l'interface de ligne de commande Hudi pour administrer les jeux de données Hudi afin d'afficher des informations sur les validations, le système de fichiers, les statistiques, etc. Vous pouvez également utiliser l'interface de ligne de commande pour effectuer manuellement des compressions, planifier des compressions ou annuler des compressions planifiées. Pour plus d'informations, consultez Interaction via la CLI dans la documentation Apache Hudi.

Pour démarrer l'interface de ligne de commande Hudi et se connecter à un ensemble de données

1. Connexion au nœud principal à l'aide de SSH Pour plus d'informations, consultez Connexion au nœud principal à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
2. Sur la ligne de commande, tapez `/usr/lib/hudi/cli/bin/hudi-cli.sh`.

   L'invite de commande devient `hudi->`.
3. Utilisez la commande suivante pour vous connecter à un ensemble de données. Remplacez *`s3://amzn-s3-demo-bucket/myhudidataset`* par le chemin d'accès à le jeu de données que vous voulez utiliser. La valeur que nous utilisons est la même que la valeur établie dans les exemples précédents.

   ```
   connect --path s3://amzn-s3-demo-bucket/myhudidataset
   ```

   L'invite de commandes change pour inclure l'ensemble de données auquel vous êtes connecté, comme illustré dans l'exemple suivant.

   ```
   hudi:myhudidataset->
   ```

Par défaut, le `hudi-cli.sh` script utilisé dans les versions 7.3.0 à 7.8.0 d'Amazon EMR est utilisé. `hudi-cli-bundle.jar` Si vous rencontrez des problèmes, vous pouvez revenir à la CLI Hudi classique à l'aide de la commande suivante :

```
/usr/lib/hudi/cli/bin/hudi-cli.sh --cliBundle false
```

Cette commande exécute le `hudi-cli.sh` script, définit l'`--cliBundle`indicateur et indique à la CLI d'utiliser les fichiers JAR individuels au lieu du JAR groupé. Par défaut, le paramètre `--cliBundle` est défini sur true, ce qui signifie que la CLI utilise plutôt le fichier JAR groupé.

Utilisation d'Amazon EMR 7.9.0 et versions supérieures

> **Note**
>
> Le script hudi-cli.sh est devenu obsolète dans les versions 7.9.0 et supérieures d'EMR. Les versions 7.9.0 et supérieures d'Amazon EMR utilisent le fichier .jar. hudi-cli-bundle

Pour démarrer l'interface de ligne de commande Hudi et se connecter à un ensemble de données :

1. Connexion au nœud principal à l'aide de SSH Pour plus d'informations, consultez Connexion au nœud principal à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
2. Sur la ligne de commande, tapez/usr/lib/hudi/cli-bundle/bin/hudi- cli-with-bundle .sh hudi-cli-with-bundleou tapez simplement >hudi-cli.

   L'invite de commande devient hudi- >.
3. Utilisez la commande suivante pour vous connecter à un ensemble de données. Remplacez s3://amzn-s3-demo-bucket/myhudidataset par le chemin d'accès à le jeu de données que vous voulez utiliser. La valeur que nous utilisons est la même que la valeur établie dans les exemples précédents.

   ```
   connect --path s3://amzn-s3-demo-bucket/myhudidataset
   ```
4. L'invite de commandes change pour inclure l'ensemble de données auquel vous êtes connecté, comme illustré dans l'exemple suivant.

   ```
   hudi:myhudidataset->
   ```

Historique des sorties de Hudi

Le tableau suivant répertorie la version de Hudi incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants

de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version de Hudi

Étiquette de version Amazon EMR	Version Hudi	Composants installés avec Hudi
emr-7.9.0	0.15.0-amzn-6	Not available.
emr-7.8.0	0.15.0-amzn-5	Not available.
emr-7,7.0	0.15.0-amzn-4	Not available.
emr-7.6.0	0.15.0-amzn-3	Not available.
emr-7.5.0	0.15.0-amzn-2	Not available.
emr-7.4.0	0.15.0-amzn-1	Not available.
emr-7.3.0	0.15.0-amzn-0	Not available.
emr-7.2.0	0.14.1-amzn-1	Not available.
emr-5.36.2	0.10.1-amzn-1	Not available.
emr-7.1.0	0.14.1-amzn-0	Not available.
emr-7.0.0	0.14.0-amzn-1	Not available.
emr-6.15.0	0.14.0-amzn-0	Not available.
emr-6.14.0	0.13.1-amzn-2	Not available.
emr-6.13.0	0.13.1-amzn-1	Not available.
emr-6.12.0	0.13.1-amzn-0	Not available.
emr-6.11.1	0.13.0-amzn-0	Not available.
emr-6.11.0	0.13.0-amzn-0	Not available.

Étiquette de version Amazon EMR	Version Hudi	Composants installés avec Hudi
emr-6.10.1	0.12.2-amzn-0	Not available.
emr-6.10.0	0.12.2-amzn-0	Not available.
emr-6.9.1	0.12.1-amzn-0	Not available.
emr-6.9.0	0.12.1-amzn-0	Not available.
emr-6.8.1	0.11.1-amzn-0	Not available.
emr-6.8.0	0.11.1-amzn-0	Not available.
emr-6.7.0	0.11.0-amzn-0	Not available.
emr-5.36.1	0.10.1-amzn-1	Not available.
emr-5.36.0	0.10.1-amzn-1	Not available.
emr-6.6.0	0.10.1-amzn-0	Not available.
emr-5.35.0	0.9.0-amzn-2	Not available.
emr-6.5.0	0.9.0-amzn-1	Not available.
emr-6.4.0	0.8.0-amzn-0	Not available.
emr-6.3.1	0.7.0-amzn-0	Not available.
emr-6.3.0	0.7.0-amzn-0	Not available.
emr-6.2.1	0.6.0-amzn-1	Not available.
emr-6.2.0	0.6.0-amzn-1	Not available.
emr-6.1.1	0.5.2-incubating-amzn-2	Not available.
emr-6.1.0	0.5.2-incubating-amzn-2	Not available.
emr-6.0.1	0.5.0-incubating-amzn-1	Not available.

Étiquette de version Amazon EMR	Version Hudi	Composants installés avec Hudi
emr-6.0.0	0.5.0-incubating-amzn-1	Not available.
emr-5.34.0	0.9.0-amzn-0	Not available.
emr-5.33.1	0.7.0-amzn-1	Not available.
emr-5.33.0	0.7.0-amzn-1	Not available.
emr-5.32.1	0.6.0-amzn-0	Not available.
emr-5.32.0	0.6.0-amzn-0	Not available.
emr-5.31.1	0.6.0-amzn-0	Not available.
emr-5.31.0	0.6.0-amzn-0	Not available.
emr-5.30.2	0.5.2-incubating	Not available.
emr-5.30.1	0.5.2-incubating	Not available.
emr-5.30.0	0.5.2-incubating	Not available.
mr-5.29.0	0.5.0-incubating	Not available.
emr-5.28.1	0.5.0-incubating	Not available.
emr-5.28.0	0.5.0-incubating	Not available.

Hue

Hue (Hadoop User Experience) est une interface utilisateur graphique open source, basée sur le web, à utiliser avec Amazon EMR et Apache Hadoop. Hue regroupe plusieurs projets d'écosystème Hadoop différents au sein d'une interface configurable. Amazon EMR a également ajouté des personnalisations spécifiques à Hue sur Amazon EMR. Hue joue le rôle de serveur frontal pour les applications exécutées sur votre cluster, ce qui vous permet d'interagir avec les applications via une interface plus conviviale ou qui vous est plus familière. Grâce aux applications figurant dans Hue, telles que les éditeurs Hive et Pig, il n'est plus nécessaire de se connecter au cluster pour exécuter des scripts de façon interactive avec le shell de chaque application. Une fois que le cluster est lancé, vous pouvez interagir entièrement avec vos applications en utilisant Hue ou une interface similaire. Pour plus d'informations sur Hue, consultez http://gethue.com.

Hue est installé par défaut lorsque vous lancez votre cluster à l'aide de la console Amazon EMR. Vous pouvez choisir de ne pas installer Hue à l'aide des Options avancées dans la console Amazon EMR lorsque vous lancez un cluster, ou en spécifiant explicitement l'option `--applications` et en omettant Hue lorsque vous utilisez `create-cluster` à partir de l' AWS CLI.

Rubriques

- Informations sur la version de Hue
- Fonctionnalités prises en charge et non prises en charge de Hue sur Amazon EMR
- Considérations
- Connexion à l'interface utilisateur Web de Hue
- Utilisation de Hue avec une base de données distante dans Amazon RDS
- Configurations avancées pour Hue
- Historique des versions de Hue

Informations sur la version de Hue

Version Hue pour 7.9.0

Le tableau suivant répertorie la version de Hue incluse dans la dernière version d’Amazon EMR série 7.x, ainsi que les composants qu’Amazon EMR installe avec Hue.

Pour connaître la version des composants installés avec Hue dans cette version, consultez Version 7.9.0.

Informations sur la version de Hue pour emr-7.9.0

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-7.9.0	Hue 4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Version de Hue pour 6.15.0

Le tableau suivant répertorie la version de Hue incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Hue.

Pour connaître la version des composants installés avec Hue dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version de Hue pour emr-6.15.0

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.15.0	Hue 4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server,

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
		mariadb-server, oozie-client, oozie-server

Version Hue pour 5.36.2

Le tableau suivant répertorie la version de Hue incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Hue.

Pour connaître la version des composants installés avec Hue dans cette version, consultez Version 5.36.2.

Informations sur la version de Hue pour emr-5.36.2

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.36.2	Hue 4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Fonctionnalités prises en charge et non prises en charge de Hue sur Amazon EMR

Avec Amazon EMR 7.0.0 et versions ultérieures, Hue requiert Python 3.9 ou version ultérieure. Lorsque vous utilisez le chiffrement en transit, vous devez utiliser une clé de chiffrement de 2 048 bits.

- Amazon S3 et Hadoop File System (HDFS) Browser
  - Avec les autorisations appropriées, vous pouvez parcourir et transférer les données entre le stockage HDFS éphémère et les compartiments S3 appartenant à votre compte.
  - Par défaut, les superutilisateurs de Hue peuvent accéder à tous les fichiers auxquels les rôles IAM Amazon EMR sont autorisés à accéder. Les utilisateurs nouvellement créés n'ont pas automatiquement les autorisations d'accéder à l'explorateur de fichiers Amazon S3 et doivent avoir les autorisations `filebrowser.s3_access` activées pour leur groupe.
- Hive – Exécutez des requêtes interactives sur vos données. Vous pouvez également créer un prototype d'interrogation par programmation ou par lot.
- Pig – Exécutez des scripts sur vos données ou exécutez des commandes interactives.
- Oozie – Créer et surveiller des flux de travail Oozie.
- Gestionnaire Metastore – Affichez et manipulez le contenu de la metastore Hive (importation/création, suppression, etc.).
- Explorateur de tâches – Affichez l'état des tâches Hadoop que vous avez soumises.
- Gestion des utilisateurs – Gérez les comptes Hue et intégrez les utilisateurs LDAP à Hue.
- AWS Exemples — Il existe plusieurs exemples « ready-to-run », qui traitent des exemples de données pour différents AWS services à l'aide des applications dans Hue. Lorsque vous vous connectez à Hue, vous êtes dirigé vers l'application Hue Home où les exemples de données sont préinstallés.
- Livy Server est pris en charge uniquement dans la version Amazon EMR 5.9.0 et ultérieure.
- Pour utiliser Hue Notebook pour Spark, vous devez installer Hue avec Livy et Spark.
- Trino/Presto – Hue prend en charge la connexion aux serveurs de données Trino et Presto. Avec Amazon EMR 7.0.0 et versions ultérieures, cette connectivité est facilitée par des connecteurs `trino-python-client` et `presto-python-client`. Les versions antérieures à 7.0.0 utilisent un connecteur JDBC via une passerelle JDBC.
- Le tableau de bord Hue n'est pas pris en charge.
- PostgreSQL n'est pas pris en charge.

Considérations

Tenez compte des limites suivantes lorsque vous utilisez Hue sur Amazon EMR.

Performances avec de grandes tables de métadonnées Hue

Si la base de données de métadonnées Hue devient trop volumineuse, les performances risquent de se dégrader. Pour vérifier la taille des tables, connectez-vous d'abord au nœud principal du EC2 cluster Amazon EMR sur Amazon avec SSH et exécutez la commande `sudo mysql -u root` pour démarrer la CLI MySQL. Pour obtenir la taille de votre table, exécutez la requête`SELECT COUNT(*) FROM hue.<table_name>`. Consultez ce qui suit pour savoir ce qui *<table_name>* peut être :

- document_bureau
- document_bureau 2
- oozie_job
- requête beeswax_saved
- séance de cire d'abeille
- histoire de beeswax_query

Si l'exécution de cette requête renvoie un nombre supérieur à 100 000, vous devez exécuter la commande de nettoyage suivante pour supprimer les anciens enregistrements.

```
cd /opt/cloudera/parcels/CDH/lib/hue # Hue home directory ./build/env/bin/hue
 desktop_document_cleanup
```

Pour plus d'informations sur le nettoyage de votre base de données, consultez la page de référence de la documentation Hue.

Hue ne nettoie pas automatiquement les tables, mais les versions 5.12.0 et supérieures d'Amazon EMR fournissent une méthode pour supprimer les anciens documents des tables. Créez le script shell suivant et exécutez-le en tant qu'étape dans un cluster Amazon EMR avec un paramètre entier qui représente le nombre maximal de jours pendant lesquels les documents doivent être conservés dans la base de données de métadonnées.

```
#!/bin/bash
if grep isMaster /mnt/var/lib/info/instance.json | grep false;
then
  echo "This is not the primary node; do nothing, exiting"
  exit 0
fi
while [ ! -f /usr/lib/hue/desktop/core/src/desktop/management/commands/
desktop_document_cleanup.py ]
```

```
do
  sleep 1
done
sudo systemctl stop hue.service
sudo sed -i 's+  LOG.warn+  # LOG.warn+g' /usr/lib/hue/desktop/core/src/desktop/
management/commands/desktop_document_cleanup.py
sudo /usr/lib/hue/build/env/bin/hue desktop_document_cleanup --keep-days $1
sudo systemctl start hue.service
```

Incompatibilité entre les versions de Hue

Si vous utilisez la même base de données de métadonnées Hue sur plusieurs clusters compatibles Hue, nous recommandons que ces clusters exécutent la même version de Hue. Les différentes versions de Hue peuvent avoir des schémas différents pour la base de données de métadonnées Hue. L'utilisation de la même base de données pour différentes versions peut entraîner l'échec d'une installation de Hue. Par exemple, essayer d'utiliser la même base de données pour deux clusters sur lesquels les versions 4.10.0 et 4.11.0 sont installées peut provoquer des erreurs de connexion pour les utilisateurs qui tentent de se connecter au cluster 4.10.0.

Pour obtenir la liste des labels de version d'Amazon EMR et des versions installées correspondantes de Hue, consultez l'historique des versions de Hue.

Connexion à l'interface utilisateur Web de Hue

La connexion à l'interface utilisateur web Hue est identique à la connexion à n'importe quelle interface HTTP hébergée sur le nœud principal d'un cluster. La procédure suivante décrit comment accéder à l'interface utilisateur de Hue. Pour plus d'informations, consultez Affichage des interfaces web hébergées sur les clusters EMR dans le Guide de gestion Amazon EMR.

Pour afficher l'interface utilisateur Web de Hue

1. Suivez ces instructions pour Configurer un tunnel SSH vers le nœud principal à l'aide du redirecteur de port dynamique dans le Guide de gestion Amazon EMR.
2. Entrez l'adresse suivante dans votre navigateur pour ouvrir l'interface Web Hue : `http://`*`master public DNS`*`:8888` où se *`master public dns`* trouve le nom DNS public de votre nœud maître de cluster, par exemple`ec2-11-22-333-44.compute-1.amazonaws.com`.
3. Sur l'écran de connexion à Hue, si vous êtes administrateur et que vous vous connectez pour la première fois, saisissez un nom d'utilisateur et un mot de passe pour créer votre supercompte

Hue, puis sélectionnez Créer un compte. Sinon, tapez votre nom d'utilisateur et votre mot de passe, puis sélectionnez Create account (Créer un compte), ou saisissez les informations d'identification fournies par votre administrateur.

Utilisation de Hue avec une base de données distante dans Amazon RDS

Par défaut, les informations utilisateur et les historiques de requêtes Hue sont stockés dans une base de données MySQL locale sur le nœud maître. Vous pouvez également créer un ou plusieurs clusters compatibles avec Hue à l'aide d'une configuration stockée dans Amazon S3 et d'une base de données MySQL dans Amazon Relational Database Service (Amazon RDS). Cela vous permet de conserver les informations utilisateur et l'historique des requêtes créés par Hue sans maintenir votre cluster Amazon EMR en cours d'exécution. Nous vous recommandons d'utiliser le chiffrement côté serveur Amazon S3 pour stocker le fichier de configuration.

Commencez par créer la base de données distante pour Hue.

Pour créer la base de données MySQL externe

1. Ouvrez la console Amazon RDS à l'adresse https://console.aws.amazon.com/rds/.
2. Cliquez sur Launch a DB Instance (Lancement d'une instance DB).
3. Choisissez MySQL et cliquez sur Select (Sélectionner).
4. Laissez la sélection par défaut de Multi-AZ Deployment and Provisioned IOPS Storage (Déploiement multi-AZ et Stockage sur volumes IOPS dimensionnés) et cliquez sur Next (Suivant).
5. Conservez les valeurs par défaut des spécifications de l'instance, spécifiez les paramètres, puis cliquez sur Next (Suivant).
6. Dans la page Configuration de paramètres avancés, choisissez un groupe de sécurité et un nom de base de données adéquats. Le groupe de sécurité que vous utilisez doit permettre au moins l'accès TCP entrant pour le port 3306 à partir du nœud maître de votre cluster. Si vous n'avez pas créé votre cluster à ce stade, vous pouvez autoriser tous les hôtes à se connecter au port 3306 et ajuster le groupe de sécurité, une fois que vous avez lancé le cluster. Cliquez sur Launch DB Instance (Lancement d'une instance DB).
7. Dans le tableau de bord RDS, sélectionnez Instances et choisissez l'instance que vous venez de créer. Lorsque votre base de données est disponible, notez la dbname (nom de la base de

données), le nom d'utilisateur, le mot de passe et le nom d'hôte d'instance RDS. Vous allez utiliser ces informations pour créer et configurer votre cluster.

Pour spécifier une base de données MySQL externe pour Hue lors du lancement d'un cluster à l'aide du AWS CLI

Pour spécifier une base de données MySQL externe pour Hue lors du lancement d'un cluster à l'aide de AWS CLI, utilisez les informations que vous avez indiquées lors de la création de votre instance RDS pour la configurer `hue.ini` avec un objet de configuration

> **Note**
>
> Vous pouvez créer plusieurs clusters qui utilisent la même base de données externe, mais chaque cluster partage l'historique des requêtes et les informations utilisateur.

- À l'aide de AWS CLI, créez un cluster sur lequel Hue est installé, en utilisant la base de données externe que vous avez créée et en référençant un fichier de configuration avec une classification de configuration pour Hue qui spécifie les propriétés de la base de données. L'exemple suivant crée un cluster avec Hue installé en référençant un fichier de configuration dans Amazon S3, `myConfig.json`, qui spécifie la configuration de la base de données.

  > **Note**
  >
  > Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

  ```
  aws emr create-cluster --release-label emr-7.9.0 --applications Name=Hue Name=Spark
   Name=Hive \
  --instance-type m5.xlarge --instance-count 3 \
  --configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/
  myConfig.json --use-default-roles
  ```

  Vous trouverez ci-dessous des exemples de contenus du fichier `myConfig.json`. Remplacez*dbname*, *usernamepassword*, et par *RDS instance hostname* les valeurs que vous avez notées précédemment dans le tableau de bord RDS.

```
[{
  "Classification": "hue-ini",
  "Properties": {},
  "Configurations": [
    {
      "Classification": "desktop",
      "Properties": {},
      "Configurations": [
        {
          "Classification": "database",
          "Properties": {
            "name": "dbname",
            "user": "username",
            "password": "password",
            "host": "RDS instance hostname",
            "port": "3306",
            "engine": "mysql"
          },
          "Configurations": []
        }
      ]
    }
  ]
}]
```

Résolution des problèmes

En cas de basculement d'Amazon RDS

Il est possible que les utilisateurs rencontrent des retards en exécutant une requête parce que l'instance de base de données Hue ne répond pas ou est en cours de basculement. Voici quelques faits et directives pour ce problème :

- Si vous vous connectez à la console Amazon RDS, vous pouvez rechercher des événements de basculement. Par exemple, pour voir si un basculement est en cours ou a eu lieu, recherchez des événements tels que « Le basculement de l'instance Multi-AZ a commencé » et « Basculement de l'instance Multi-AZ terminé. »
- Il faut environ 30 secondes à une instance RDS pour terminer un basculement.

- Si vous recevez des longer-than-normal réponses pour des requêtes dans Hue, essayez de réexécuter la requête.

Configurations avancées pour Hue

Cette section comprend les rubriques suivantes.

Rubriques

- Configuration de Hue pour les utilisateurs LDAP

Configuration de Hue pour les utilisateurs LDAP

L'intégration sur LDAP permet aux utilisateurs de se connecter à Hue à l'aide d'informations d'identification existantes stockées dans un annuaire LDAP. Lorsque vous intégrez Hue sur LDAP, vous n'avez pas besoin de gérer indépendamment des informations utilisateur dans Hue. Les informations ci-dessous illustrent l'intégration de Hue avec Microsoft Active Directory, mais les options de configuration sont analogues à n'importe quel répertoire LDAP.

Tout d'abord, l'authentification LDAP se lie au serveur et établit la connexion. Ensuite, la connexion établie est utilisée pour toutes les requêtes suivantes pour rechercher des informations utilisateur LDAP. A moins que votre serveur Active Directory autorise les connexions anonymes, une connexion doit être établie à l'aide d'un nom unique et d'un mot de passe de liaison. Le nom unique (ou DN) de liaison est défini par le paramètre de configuration `bind_dn`. Le mot de passe de liaison est défini par le paramètre de configuration `bind_password`. Hue a deux manières de lier des demandes LDAP : liaison de recherche et liaison directe. La méthode préférée pour l'utilisation de Hue avec Amazon EMR est la liaison de recherche.

Lorsque la liaison de recherche est utilisée avec Active Directory, Hue utilise l'attribut de nom d'utilisateur (défini par `user_name_attr config`) pour rechercher l'attribut qui doit être extrait du nom unique (ou DN) de la base. La liaison de recherche est utile lorsque le nom unique complet n'est pas connu pour l'utilisateur de Hue.

Par exemple, vous avez peut-être `user_name_attr config` définir pour utiliser le nom commun (ou CN). Dans ce cas, le serveur Active Directory utilise le nom d'utilisateur Hue fourni au cours de l'ouverture de session pour rechercher dans l'arborescence de répertoires un nom commun qui corresponde, en commençant par le nom unique de la base. Si le nom commun pour l'utilisateur Hue

est trouvé, le nom unique de l'utilisateur est renvoyé par le serveur. Hue construit ensuite un nom unique utilisé pour authentifier l'utilisateur en effectuant une opération de liaison.

> **Note**
>
> La liaison de recherche parcourt toutes les sous-arborescences du répertoire pour trouver des noms d'utilisateur, en commençant par le nom unique de la base. Le nom unique de la base spécifié dans la configuration LDAP Hue doit être le plus proche parent du nom d'utilisateur, à défaut de quoi les performances de votre authentification LDAP pourront en être affectées.

Lorsque la liaison directe est utilisée avec Active Directory, le `nt_domain` ou `ldap_username_pattern` exact doit être utilisé pour s'authentifier. Lorsque la liaison directe est utilisée, si l'attribut de domaine nt (défini par le paramètre de configuration `nt_domain`) est défini, un modèle de nom unique d'utilisateur est créé à l'aide du formulaire : `<login username>@nt_domain`. Ce modèle est utilisé pour rechercher dans toutes les sous-arborescences du répertoire, en commençant par le nom unique de la base. Si le domaine nt n'est pas configuré, Hue recherche un modèle exact de nom unique pour l'utilisateur (défini par le paramètre de configuration `ldap_username_pattern`). Dans ce cas, le serveur recherche une valeur `ldap_username_pattern` correspondante dans toutes les sous-arborescences du répertoire, en commençant par le nom unique de la base.

Pour lancer un cluster avec des propriétés LDAP pour Hue à l'aide de l' AWS CLI

- Pour spécifier les propriétés LDAP pour `hue-ini`, créez un cluster avec Hue installé et référencez un fichier JSON avec les propriétés de configuration pour LDAP. Un exemple de commande est présenté ci-dessous. Elle fait référence à un fichier de configuration `myConfig.json` stocké dans Amazon S3.

  ```
  aws emr create-cluster --release-label emr-7.9.0 --applications Name=Hue Name=Spark
   Name=Hive \
  --instance-type m5.xlarge --instance-count 3 --configurations https://
  s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json.
  ```

  Vous trouverez ci-dessous des exemples de contenus `myConfig.json`.

  ```
  [
      {
  ```

```
        "Classification": "hue-ini",
        "Properties": {},
        "Configurations": [
            {
                "Classification": "desktop",
                "Properties": {},
                "Configurations": [
                    {
                        "Classification": "ldap",
                        "Properties": {},
                        "Configurations": [
                            {
                                "Classification": "ldap_servers",
                                "Properties": {},
                                "Configurations": [
                                    {
                                        "Classification": "yourcompany",
                                        "Properties": {
                                            "base_dn":
"DC=yourcompany,DC=hue,DC=com",
                                            "ldap_url": "ldap://ldapurl",
                                            "search_bind_authentication": "true",
                                            "bind_dn":
"CN=hue,CN=users,DC=yourcompany,DC=hue,DC=com",
                                            "bind_password": "password"
                                        },
                                        "Configurations": []
                                    }
                                ]
                            }
                        ]
                    },
                    {
                        "Classification": "auth",
                        "Properties": {
                         "backend": "desktop.auth.backend.LdapBackend"
                              }
                    }
                ]
            }

        ]
    }
```

```
]
```

> **Note**
>
> Avec la version 5.21.0 et ultérieures d'Amazon EMR, vous permet de remplacer les configurations de cluster et de spécifier des classifications de configuration supplémentaires pour chaque groupe d'instances dans un cluster en cours d'exécution. Pour ce faire, utilisez la console Amazon EMR, le AWS Command Line Interface (AWS CLI) ou le AWS SDK. Pour plus d'informations, consultez Fourniture d'une configuration pour un groupe d'instances dans un cluster en cours d'exécution.

Pour afficher des paramètres LDAP dans Hue

1. Vérifiez que vous disposez d'une connexion VPN active ou d'un tunnel SSH pour le nœud principal du cluster Amazon EMR. Tapez ensuite : 8888 dans votre navigateur *master-public-dns* pour ouvrir l'interface Web de Hue.
2. Connectez-vous à l'aide de vos informations d'identification d'administrateur Hue. Si la fenêtre Did you know? (Le saviez vous ?) s'ouvre, cliquez sur Got it, prof! (Compris, professeur !) pour la fermer.
3. Cliquez sur l'icône Hue dans la barre d'outils.
4. Sur la page About Hue (À propos de Hue), sélectionnez Configuration.
5. Dans la section Configuration Sections and Variables (Sections et variables de configuration), cliquez sur Desktop (Bureau).
6. Défilez à la section ldap pour afficher vos paramètres.

Historique des versions de Hue

Le tableau suivant répertorie la version de Hue incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version de Hue

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-7.9.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.8.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7,7.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server,

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
		mariadb-server, oozie-client, oozie-server
emr-7.6.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.5.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-7.4.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.3.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-7.2.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.36.2	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-7.1.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.0.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.15.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.14.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.13.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.12.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.11.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.11.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.10.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.10.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.9.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.9.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.8.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.8.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.7.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.36.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.36.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.6.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.35.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.5.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.4.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.3.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.3.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.2.1	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.2.0	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.1.1	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.1.0	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.0.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-6.0.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.34.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.33.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.33.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.32.1	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.32.0	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.31.1	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.31.0	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.30.2	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.30.1	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.30.0	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
mr-5.29.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.28.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.28.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.27.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.27.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.26.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.25.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.24.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.24.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.23.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.23.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.22.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.21.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.21.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.21.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.20.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.20.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.19.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.19.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.18.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.18.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.17.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.17.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.17.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.16.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.16.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.15.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.15.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.14.2	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.14.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.14.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.13.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.13.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.12.3	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.12.2	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.12.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.12.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.11.4	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.3	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.11.2	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.11.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.10.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.10.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.9.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.9.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.8.3	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.8.2	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.8.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.8.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.7.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.7.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.6.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.6.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.4	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.3	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.5.2	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.4.1	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.4.0	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.3.2	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.3.1	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.3.0	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.3	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.2.2	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.1.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.1.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.0.3	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-5.0.2	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-5.0.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-5.0.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-4.9.6	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.5	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-4.9.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-4.8.5	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-4.8.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-4.7.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.7.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.7.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-4.7.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.7.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.6.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-4.6.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.5.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.4.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Étiquette de version Amazon EMR	Version Hue	Composants installés avec Hue
emr-4.3.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.2.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.1.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Iceberg

Apache Iceberg est un format de tableau ouvert pour les grands jeux de données dans Amazon Simple Storage Service (Amazon S3). Il fournit des performances de requête rapides sur de grandes tables, des validations atomiques, des écritures simultanées et une évolution de table compatible avec SQL. À partir d'Amazon EMR 6.5.0, vous pouvez utiliser Apache Spark 3 sur des clusters Amazon EMR au format de table Iceberg.

Le tableau suivant répertorie la version d'Iceberg incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants qu'Amazon EMR installe avec Iceberg.

Pour connaître la version des composants installés avec Iceberg dans cette version, consultez la version 7.9.0 Versions des composants.

Informations sur la version d'Iceberg pour emr-7.9.0

Étiquette de version Amazon EMR	Version Iceberg	Composants installés avec Iceberg
emr-7.9.0	Iceberg 1.7.1-amzn-2	Not available.

Le tableau suivant répertorie la version d'Iceberg incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Iceberg.

Pour connaître la version des composants installés avec Iceberg dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version d'Iceberg pour emr-6.15.0

Étiquette de version Amazon EMR	Version Iceberg	Composants installés avec Iceberg
emr-6.15.0	Iceberg 1.4.0-amzn-0	Not available.

Rubriques

- Comment fonctionne Iceberg
- Utiliser un cluster avec Iceberg

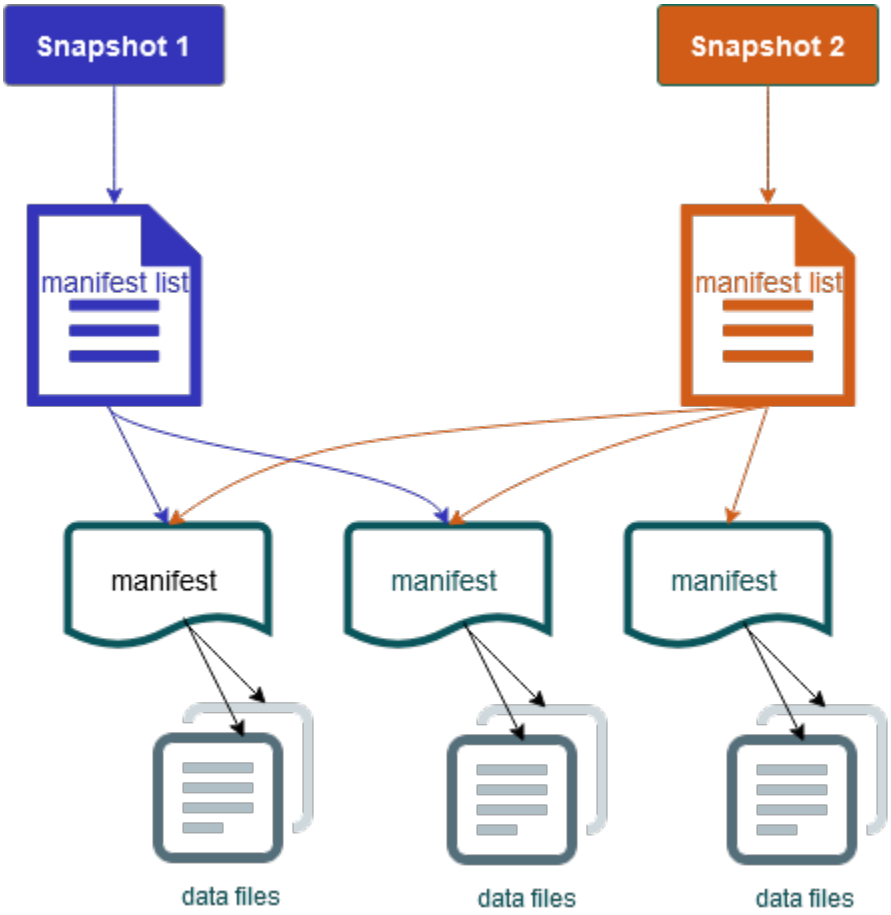
- Historique des versions d'Iceberg

Comment fonctionne Iceberg

Iceberg suit les fichiers de données individuels dans une table plutôt que dans des répertoires. Ainsi, les rédacteurs peuvent créer des fichiers de données sur place (les fichiers ne sont ni déplacés ni modifiés). De plus, les rédacteurs ne peuvent ajouter des fichiers à la table que dans le cadre d'un commit explicite. L'état de la table est conservé dans les fichiers de métadonnées. Toutes les modifications apportées à l'état de la table créent un nouveau fichier de métadonnées qui remplace automatiquement les anciennes métadonnées. Le fichier de métadonnées de table suit le schéma de la table, la configuration du partitionnement et d'autres propriétés.

Il inclut également des instantanés du contenu du tableau. Chaque instantané est un ensemble complet de fichiers de données figurant dans le tableau à un moment donné. Les instantanés sont répertoriés dans le fichier de métadonnées, mais les fichiers d'un instantané sont stockés dans des fichiers manifestes distincts. Les transitions atomiques d'un fichier de métadonnées de table à l'autre permettent d'isoler les instantanés. Les lecteurs utilisent l'instantané actuel lorsqu'ils ont chargé les métadonnées de la table. Les lecteurs ne sont pas affectés par les modifications tant qu'ils n'ont pas actualisé et sélectionné un nouvel emplacement de métadonnées. Les fichiers de données des instantanés sont stockés dans un ou plusieurs fichiers manifestes qui contiennent une ligne pour chaque fichier de données de la table, ses données de partition et ses métriques. Un instantané est l'union de tous les fichiers contenus dans ses manifestes. Les fichiers manifestes peuvent également être partagés entre les instantanés afin d'éviter de réécrire des métadonnées peu modifiées.

Schéma d'instantané d'Iceberg



Iceberg offre les fonctionnalités suivantes :

- Prend en charge les transactions ACID et les voyages dans le temps dans votre lac de données Amazon S3.
- Les nouvelles tentatives de validation bénéficient des avantages en termes de performances d'une simultanéité optimiste.
- La résolution des conflits au niveau des fichiers entraîne une forte simultanéité.
- Avec des statistiques min-max par colonne dans les métadonnées, vous pouvez ignorer des fichiers, ce qui améliore les performances des requêtes sélectives.
- Vous pouvez organiser les tables selon des configurations de partition flexibles, l'évolution des partitions permettant de mettre à jour les schémas de partition. Les requêtes et les volumes de données peuvent alors changer sans recourir à des répertoires physiques.
- Prend en charge l'évolution des schémas et leur application.
- Les tables Iceberg agissent comme des puits idempotents et des sources rejouables. Cela permet le streaming et le support par lots avec des pipelines uniques. Les récepteurs idempotents suivent

les opérations d'écriture réussies dans le passé. Par conséquent, le récepteur peut à nouveau demander des données en cas de panne et supprimer des données si elles ont été envoyées plusieurs fois.

- Consultez l'historique et le lignage, y compris l'évolution des tables, l'historique des opérations et les statistiques pour chaque validation.
- Migrez depuis un jeu de données existant avec un choix de format de données (Parquet, ORC, Avro) et de moteur d'analyse (Spark, Trino, PrestoDB, Flink, Hive).

Utiliser un cluster avec Iceberg

Cette section contient des informations sur l'utilisation d'Iceberg avec Spark, Trino, Flink et Hive.

Utiliser un cluster Iceberg avec Spark

À partir de la version 6.5.0 d'Amazon EMR, vous pouvez utiliser Iceberg avec votre cluster Spark sans avoir à inclure d'actions de démarrage. Pour les versions 6.4.0 et antérieures d'Amazon EMR, vous pouvez utiliser une action d'amorçage pour préinstaller toutes les dépendances nécessaires.

Dans ce didacticiel, vous allez utiliser le AWS CLI pour travailler avec Iceberg sur un cluster Amazon EMR Spark. Pour utiliser la console afin de créer un cluster avec Iceberg installé, suivez les étapes de la section Création d'un lac de données Apache Iceberg à l'aide d'Amazon Athena, d'Amazon EMR et d' AWS Glue.

Création d'un cluster Iceberg

Vous pouvez créer un cluster avec Iceberg installé à l'aide de l' AWS Management Console API, AWS CLI ou de l'API Amazon EMR. Dans ce didacticiel, vous allez utiliser le AWS CLI pour travailler avec Iceberg sur un cluster Amazon EMR. Pour utiliser la console afin de créer un cluster avec Iceberg installé, suivez les étapes de la section Création d'un lac de données Apache Iceberg à l'aide d'Amazon Athena, d'Amazon EMR et d' AWS Glue.

Pour utiliser Iceberg sur Amazon EMR avec AWS CLI le, créez d'abord un cluster en suivant les étapes suivantes. Pour plus d'informations sur la spécification de la classification des icebergs à l'aide du AWS CLI, voir Fournir une configuration à l'aide de l' AWS CLI option lorsque vous créez un cluster ouFournir une configuration à l'aide du kit SDK Java lors de la création d'un cluster.

1. Créez un fichier `configurations.json` avec le contenu suivant :

```
[{
    "Classification":"iceberg-defaults",
    "Properties":{"iceberg.enabled":"true"}
}]
```

2. Créez ensuite un cluster à l'aide de la configuration suivante. Remplacez l'exemple de chemin de compartiment Amazon S3 et l'ID de sous-réseau par les vôtres.

```
aws emr create-cluster --release-label emr-6.5.0 \
--applications Name=Spark \
--configurations file://configurations.json \
--region us-east-1 \
--name My_Spark_Iceberg_Cluster \
--log-uri s3://amzn-s3-demo-bucket/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Vous pouvez également créer un cluster Amazon EMR incluant l'application Spark et inclure le fichier `/usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar` en tant que dépendance JAR dans une tâche Spark. Pour plus d'informations, consultez Soumission d'applications.

Pour inclure le fichier jar en tant que dépendance dans une tâche Spark, ajoutez la propriété de configuration suivante à l'application Spark :

```
--conf "spark.jars=/usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar"
```

Pour plus d'informations sur les dépendances des tâches Spark, consultez la section Gestion des dépendances dans le document d'Apache Spark Running Spark on Kubernetes.

Initialisation d'une session Spark pour Iceberg

Les exemples suivants montrent comment lancer le shell Spark interactif, utiliser Spark submit ou utiliser Blocs-notes EMR pour travailler avec Iceberg sur Amazon EMR.

spark-shell

1. Connexion au nœud principal à l'aide de SSH Pour plus d'informations, consultez Connexion au nœud principal à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
2. Entrez la commande suivante pour lancer le shell Spark. Pour utiliser la PySpark coque, remplacez-la `spark-shell` par`pyspark`.

```
spark-shell \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket/prefix/
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

spark-submit

1. Connexion au nœud principal à l'aide de SSH Pour plus d'informations, consultez Connexion au nœud principal à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
2. Entrez la commande suivante pour lancer la session Spark pour Iceberg.

```
spark-submit \
--conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
--conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
--conf spark.sql.catalog.my_catalog.type=glue \
--conf spark.sql.defaultCatalog=my_catalog \
--conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

EMR Studio notebooks

Pour initialiser une session Spark à l'aide des blocs-notes EMR Studio, configurez votre session Spark à l'aide de la commande magique `%%configure` de votre bloc-notes Amazon EMR, comme dans l'exemple suivant. Pour plus d'informations, consultez Utilisation des magies de Blocs-notes EMR dans le Guide de gestion Amazon EMR.

```
%%configure -f{
"conf":{
    "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
```

```
    "spark.sql.catalog.my_catalog.type":"glue",
    "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
    "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
}
```

CLI

Pour initialiser un cluster Spark à l'aide de la CLI et définir toutes les configurations par défaut de session Spark Iceberg, exécutez l'exemple suivant. Pour plus d'informations sur la spécification d'une classification de configuration à l'aide de l'API Amazon EMR AWS CLI et Amazon EMR, consultez Configurer les applications.

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
      "spark.sql.catalog.my_catalog.type":"glue",
      "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
      "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
  }
]
```

Écriture dans une table Iceberg

L'exemple suivant montre comment créer un jeu de données Iceberg DataFrame et l'écrire sous forme de jeu de données. Les exemples illustrent l'utilisation des jeux de données à l'aide du shell Spark lorsque vous êtes connecté au nœud principal à l'aide de SSH comme utilisateur hadoop par défaut.

> **Note**
>
> Pour coller des exemples de code dans le shell Spark, tapez `:paste` à l'invite, collez l'exemple, puis appuyez sur `CTRL+D`.

PySpark

Spark inclut un shell basé sur Python, `pyspark`, que vous pouvez utiliser pour créer un prototype de programmes Spark écrits en Python. Invoquez `pyspark` sur le nœud principal.

```
## Create a DataFrame.
data = spark.createDataFrame([
 ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
 ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
 ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
 ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z")
],["id", "creation_date", "last_update_time"])

## Write a DataFrame as a Iceberg dataset to the Amazon S3 location.
spark.sql("""CREATE TABLE IF NOT EXISTS dev.db.iceberg_table (id string,
creation_date string,
last_update_time string)
USING iceberg
location 's3://amzn-s3-demo-bucket/example-prefix/db/iceberg_table'""")

data.writeTo("dev.db.iceberg_table").append()
```

Scala

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._

// Create a DataFrame.
val data = Seq(
("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103", "2015-01-01", "2015-01-01T13:51:40.519832Z")
).toDF("id", "creation_date", "last_update_time")

// Write a DataFrame as a Iceberg dataset to the Amazon S3 location.
spark.sql("""CREATE TABLE IF NOT EXISTS dev.db.iceberg_table (id string,
creation_date string,
last_update_time string)
USING iceberg
location 's3://amzn-s3-demo-bucket/example-prefix/db/iceberg_table'""")

data.writeTo("dev.db.iceberg_table").append()
```

Lecture à partir d'une table Iceberg

PySpark

```
df = spark.read.format("iceberg").load("dev.db.iceberg_table")
df.show()
```

Scala

```
val df = spark.read.format("iceberg").load("dev.db.iceberg_table")
df.show()
```

Spark SQL

```
SELECT * from dev.db.iceberg_table LIMIT 10
```

Utilisation du catalogue AWS de données Glue avec Spark Iceberg

Vous pouvez vous connecter à AWS Glue Data Catalog depuis Spark Iceberg. Cette section présente les différentes commandes de connexion.

Connectez-vous au catalogue AWS Glue par défaut dans votre région par défaut

Cet exemple montre comment se connecter à l'aide du type de catalogue Glue. Si vous ne spécifiez pas d'ID de catalogue, il utilise la valeur par défaut :

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Connectez-vous à un catalogue AWS Glue avec un identifiant de catalogue spécifique

Cet exemple montre comment se connecter à l'aide d'un identifiant de catalogue :

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
```

```
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.catalog.my_catalog.glue.id=AWS Glue catalog ID \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Cette commande peut être utilisée pour se connecter à un catalogue AWS Glue dans un autre compte, à un catalogue RMS ou à un catalogue fédéré.

Utilisation du catalogue REST (IRC) d'Iceberg avec Spark Iceberg

Les sections suivantes expliquent comment configurer l'intégration d'Iceberg à un catalogue.

Connect to AWS Glue Data Catalog IRC endpoint

Voici un exemple de `spark-submit` commande pour utiliser Iceberg REST :

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=glue catalog ID \
    --conf spark.sql.catalog.my_catalog.type=rest \
    --conf spark.sql.catalog.my_catalog.uri=glue endpoint URI/iceberg \
    --conf spark.sql.catalog.my_catalog.rest.sigv4-enabled=true \
    --conf spark.sql.catalog.my_catalog.rest.signing-name=glue \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Pour l'utiliser sur un cluster activé pour les rôles d'exécution, les paramètres de configuration Spark supplémentaires suivants sont nécessaires :

```
"spark.hadoop.fs.s3.credentialsResolverClass":
 "software.amazon.glue.GlueTableCredentialsResolver",
"spark.hadoop.catalog-impl": "org.apache.iceberg.aws.glue.GlueCatalog",
"spark.hadoop.glue.id": glue catalog ID
"spark.hadoop.glue.endpoint": "glue endpoint"
```

Pour AWS la liste des URL des points de terminaison Glue pour chaque région, consultez la section Points de terminaison et quotas de AWS Glue.

Connectez-vous à un point de terminaison IRC arbitraire

Voici un exemple de `spark-submit` commande pour utiliser un point de terminaison IRC :

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=warehouse name \
    --conf spark.sql.catalog.my_catalog.type=rest \
    --conf spark.sql.catalog.my_catalog.uri=your rest endpoint \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Différences de configuration entre Iceberg SparkCatalog et SparkSessionCatalog

Iceberg propose deux méthodes pour créer des catalogues Spark Iceberg. Vous pouvez définir la configuration de Spark sur l'`SparkCatalog`une ou sur`SparkSessionCatalog`.

Utiliser Iceberg SparkCatalog

Voici la commande à utiliser SparkCatalogcomme catalogue Spark Iceberg :

```
spark-shell \
--conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
--conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
--conf spark.sql.catalog.my_catalog.type=glue \
--conf spark.sql.defaultCatalog=my_catalog
```

Considérations relatives à cette approche :

- Vous pouvez accéder aux tables Iceberg, mais pas aux autres tables.
- Le nom du catalogue ne peut pas être spark_catalog. Il s'agit du nom du catalogue initial dans Spark. Il se connecte toujours à un métastore Hive. Il s'agit du catalogue par défaut dans Spark, sauf si l'utilisateur le remplace à l'aide `spark.sql.defaultCatalog` de.
- Vous pouvez définir le nom `spark.sql.defaultCatalog` de votre catalogue pour en faire le catalogue par défaut.

Utiliser Iceberg SparkSessionCatalog

Voici la commande à utiliser SparkSessionCatalogcomme catalogue Spark Iceberg :

```
spark-shell \
    --conf spark.sql.catalog.spark_catalog=org.apache.iceberg.spark.SparkSessionCatalog
 \
```

```
    --conf spark.sql.catalog.spark_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
    --conf spark.sql.catalog.spark_catalog.type=glue
```

Considérations relatives à cette approche :

- Si aucune table n'est trouvée sous forme de table Iceberg, Spark essaiera de voir s'il s'agit d'une table dans le métastore Hive. Voir Utiliser le catalogue de données AWS Glue comme catalogue pour Hive pour plus d'informations.
- Le nom du catalogue doit être spark_catalog.

Utilisation des extensions Iceberg Spark

Iceberg propose une extension Spark `org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions` que les utilisateurs peuvent définir via la configuration `spark.sql.extensions` des extensions Spark. Les extensions activent les fonctionnalités clés d'Iceberg, telles que DELETE, UPDATE et MERGE au niveau des lignes, les instructions et procédures du langage de définition des données Spark spécifiques à Iceberg, telles que le compactage, l'expiration des instantanés, le branchement et le balisage, etc. Consultez ce qui suit pour plus de détails :

- Extensions d'écriture d'Iceberg Spark : Spark Writes
- Extensions DDL Iceberg Spark : extensions SQL ALTER TABLE
- Extensions de procédure Iceberg Spark : procédures Spark

Considérations relatives à l'utilisation d'Iceberg avec Spark

- Amazon EMR 6.5.0 ne prend pas en charge l'exécution d'Iceberg sur Amazon EMR sur EKS par défaut. Une image personnalisée Amazon EMR 6.5.0 est disponible afin que vous puissiez la transmettre `--jars local:///usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar` en tant que paramètre `spark-submit` pour créer des tables Iceberg sur Amazon EMR sur EKS. Pour plus d'informations, consultez Soumettre une charge de travail Spark dans Amazon EMR à l'aide d'une image personnalisée dans le Guide de développement Amazon EMR sur EKS. Vous pouvez également nous contacter Support pour obtenir de l'aide. À partir d'Amazon EMR 6.6.0, Iceberg est pris en charge sur Amazon EMR sur EKS.
- Lorsque vous utilisez AWS Glue comme catalogue pour Iceberg, assurez-vous que la base de données dans laquelle vous créez une table existe dans AWS Glue. Si vous utilisez des services

tels que AWS Lake Formation et que vous ne parvenez pas à charger le catalogue, assurez-vous de disposer d'un accès approprié au service pour exécuter la commande.

- Si vous utilisez Iceberg SparkSessionCatalog, comme décrit dansDifférences de configuration entre Iceberg SparkCatalog et SparkSessionCatalog, vous devez suivre les étapes de configuration décrites dans Configurer le catalogue de données AWS Glue en tant que métastore Apache Hive, en plus de configurer les paramètres du catalogue de données Spark Iceberg Glue AWS .

Utiliser un cluster Iceberg avec Trino

À partir de la version 6.6.0 d'Amazon EMR, vous pouvez utiliser Iceberg avec votre cluster Trino.

Dans ce didacticiel, vous allez utiliser le AWS CLI pour travailler avec Iceberg sur un cluster Amazon EMR Trino. Pour utiliser la console afin de créer un cluster avec Iceberg installé, suivez les étapes de la section Création d'un lac de données Apache Iceberg à l'aide d'Amazon Athena, d'Amazon EMR et d' AWS Glue.

Création d'un cluster Iceberg

Pour utiliser Iceberg sur Amazon EMR avec AWS CLI le, créez d'abord un cluster en suivant les étapes suivantes. Pour plus d'informations sur la spécification de la classification des icebergs à l'aide du AWS CLI, voir Fournir une configuration à l'aide de l' AWS CLI option lorsque vous créez un cluster ouFournir une configuration à l'aide du kit SDK Java lors de la création d'un cluster.

1. Créez un fichier `iceberg.properties` et définissez une valeur pour le catalogue de votre choix. Par exemple, si vous souhaitez utiliser le métastore Hive comme catalogue, le contenu de votre fichier doit être le suivant.

   ```
   connector.name=iceberg
   hive.metastore.uri=thrift://localhost:9083
   ```

   Si vous souhaitez utiliser le AWS Glue Data Catalog comme boutique, le contenu de votre fichier doit être le suivant.

   ```
   connector.name=iceberg
   iceberg.catalog.type=glue
   ```

2. Créez une action d'amorçage qui copiée `iceberg.properties` depuis Amazon S3 vers `/etc/trino/conf/catalog/iceberg.properties`, comme dans l'exemple suivant.

Pour plus d'informations sur les actions d'amorçage, consultez la section Création d'actions d'amorçage pour installer des logiciels supplémentaires.

```
set -ex
sudo aws s3 cp s3://amzn-s3-demo-bucket/iceberg.properties /etc/trino/conf/catalog/
iceberg.properties
```

3. Créez un cluster avec la configuration suivante, en remplaçant l'exemple de chemin du script d'actions amorçage et le nom de clé par les vôtres.

```
aws emr create-cluster --release-label emr-6.7.0 \
--applications Name=Trino \
--region us-east-1 \
--name My_Trino_Iceberg_Cluster \
--bootstrap-actions '[{"Path":"s3://amzn-s3-demo-bucket","Name":"Add
 iceberg.properties"}]' \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=c3.4xlarge
 InstanceGroupType=CORE,InstanceCount=3,InstanceType=c3.4xlarge \
--use-default-roles \
--ec2-attributes KeyName=<key-name>
```

Initialisation d'une session Trino pour Iceberg

Pour initialiser une session Trino, exécutez la commande suivante.

```
trino-cli --catalog iceberg
```

Écriture dans une table Iceberg

Créez et écrivez dans votre table à l'aide des commandes SQL suivantes.

```
trino> SHOW SCHEMAS;
trino> CREATE TABLE default.iceberg_table (
            id int,
            data varchar,
            category varchar)
       WITH (
            format = 'PARQUET',
            partitioning = ARRAY['category', 'bucket(id, 16)'],
            location = 's3://amzn-s3-demo-bucket/<prefix>')
```

```
trino> INSERT INTO default.iceberg_table VALUES (1,'a','c1'), (2,'b','c2'),
 (3,'c','c3');
```

Lire à partir d'une table pour Iceberg

Pour lire les données de votre table Iceberg, exécutez la commande suivante.

```
trino> SELECT * from default.iceberg_table;
```

Considérations relatives à l'utilisation d'Iceberg avec Trino

- Amazon EMR 6.5 ne prend pas en charge le catalogue Trino Iceberg pour Iceberg en mode natif. Trino a besoin d'Iceberg v0.11. Nous vous recommandons donc de lancer un cluster Amazon EMR pour Trino distinct du cluster Spark et d'inclure Iceberg v0.11 sur ce cluster.
- Lorsque vous utilisez AWS Glue comme catalogue pour Iceberg, assurez-vous que la base de données dans laquelle vous créez une table existe dans AWS Glue. Si vous utilisez des services tels que AWS Lake Formation et que vous ne parvenez pas à charger le catalogue, assurez-vous de disposer d'un accès approprié au service pour exécuter la commande.
- L'intégration d'Iceberg Glue ne fonctionne pas avec le catalogue Redshift Managed Storage.

Utiliser un cluster Iceberg avec Flink

À partir de la version 6.9.0 d'Amazon EMR, vous pouvez utiliser Iceberg avec un cluster Flink sans avoir à effectuer les étapes de configuration requises lors de l'utilisation de l'intégration open source Iceberg Flink.

Création d'un cluster Iceberg

Vous pouvez créer un cluster avec Iceberg installé en utilisant la AWS Management Console, l' AWS CLI ou l'API Amazon EMR. Dans ce didacticiel, vous allez utiliser le AWS CLI pour travailler avec Iceberg sur un cluster Amazon EMR. Pour utiliser la console afin de créer un cluster avec Iceberg installé, suivez les étapes de la section Création d'un lac de données Apache Iceberg à l'aide d'Amazon Athena, d'Amazon EMR et d' AWS Glue.

Pour utiliser Iceberg sur Amazon EMR avec AWS CLI le, créez d'abord un cluster en suivant les étapes suivantes. Pour plus d'informations sur la spécification de la classification des icebergs à l'aide du AWS CLI, voir Fournir une configuration à l'aide de l' AWS CLI option lorsque vous créez un

cluster ouFournir une configuration à l'aide du kit SDK Java lors de la création d'un cluster. Créez un fichier nommé `configurations.json` avec le contenu suivant :

```
[{
"Classification":"iceberg-defaults",
    "Properties":{"iceberg.enabled":"true"}
}]
```

Créez ensuite un cluster avec la configuration suivante, en remplaçant les exemples de chemin de compartiment et d'ID de sous-réseau Amazon S3 par vos propres valeurs :

```
aws emr create-cluster --release-label emr-6.9.0 \
--applications Name=Flink \
--configurations file://iceberg_configurations.json \
--region us-east-1 \
--name My_flink_Iceberg_Cluster \
--log-uri s3://amzn-s3-demo-bucket/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef
```

Vous pouvez également créer un cluster Amazon EMR 6.9.0 contenant une application Flink et utiliser le fichier `/usr/share/aws/iceberg/lib/iceberg-flink-runtime.jar` comme dépendance JAR dans une tâche Flink.

Utilisation du client Flink SQL

Le script du client SQL se trouve sous `/usr/lib/flink/bin`. Vous pouvez exécuter le script à l'aide de la commande suivante :

```
flink-yarn-session -d # starting the Flink YARN Session in detached mode
./sql-client.sh
```

Cela lance un Flink SQL Shell.

Exemples de Flink

Création d'une table Iceberg

Flink SQL

```
CREATE CATALOG glue_catalog WITH (
   'type'='iceberg',
   'warehouse'='<WAREHOUSE>',
   'catalog-type'='glue'
 );

USE CATALOG  glue_catalog;

CREATE DATABASE IF NOT EXISTS <DB>;

USE <DB>;

CREATE TABLE IF NOT EXISTS `glue_catalog`.`<DB>`.`sample` (id int, data string);
```

API de table

```
EnvironmentSettings settings =
                EnvironmentSettings.newInstance().inBatchMode().build();

TableEnvironment tEnv = TableEnvironment.create(settings);

String warehouse = "<WAREHOUSE>";
String db = "<DB>";

tEnv.executeSql(
                "CREATE CATALOG glue_catalog WITH (\n"
                        + "   'type'='iceberg',\n"
                        + "   'warehouse'='"
                        + warehouse
                        + "',\n"
                        + "   'catalog-impl'='org.apache.iceberg.aws.glue.GlueCatalog',
\n"
                        + "   'io-impl'='org.apache.iceberg.aws.s3.S3FileIO'\n"
                        + " );");

tEnv.executeSql("USE CATALOG  glue_catalog;");
tEnv.executeSql("CREATE DATABASE IF NOT EXISTS " + db + ";");
tEnv.executeSql("USE " + db + ";");
tEnv.executeSql(
        "CREATE TABLE `glue_catalog`.`" + db + "`.`sample` (id bigint, data string);");
```

Écriture dans une table Iceberg

Flink SQL

```
INSERT INTO `glue_catalog`.`<DB>`.`sample` values (1, 'a'),(2,'b'),(3,'c');
```

API de table

```
tEnv.executeSql(
        "INSERT INTO `glue_catalog`.`"
                + db
                + "`.`sample` values (1, 'a'),(2,'b'),(3,'c');");
```

API Datastream

```
final StreamExecutionEnvironment env =
 StreamExecutionEnvironment.getExecutionEnvironment();

final StreamTableEnvironment tableEnv = StreamTableEnvironment.create(env);

String db = "<DB Name>";

String warehouse = "<Warehouse Path>";

GenericRowData rowData1 = new GenericRowData(2);
rowData1.setField(0, 1L);
rowData1.setField(1, StringData.fromString("a"));

DataStream<RowData> input = env.fromElements(rowData1);

Map<String, String> props = new HashMap<();
props.put("type", "iceberg");
props.put("warehouse", warehouse);
props.put("io-impl", "org.apache.iceberg.aws.s3.S3FileIO");

CatalogLoader glueCatlogLoader =
        CatalogLoader.custom(
                "glue",
                props,
                new Configuration(),
                "org.apache.iceberg.aws.glue.GlueCatalog");

TableLoader tableLoader =
```

```
        TableLoader.fromCatalog(glueCatlogLoader, TableIdentifier.of(db, "sample"));

DataStreamSink<Void> dataStreamSink =
        FlinkSink.forRowData(input).tableLoader(tableLoader).append();

env.execute("Datastream Write");
```

Lecture à partir d'une table Iceberg

Flink SQL

```
SELECT * FROM `glue_catalog`.`<DB>`.`sample`;
```

API de table

```
Table result = tEnv.sqlQuery("select * from `glue_catalog`.`" + db + "`.`sample`;");
```

API Datastream

```
final StreamExecutionEnvironment env =
 StreamExecutionEnvironment.getExecutionEnvironment();

final StreamTableEnvironment tableEnv = StreamTableEnvironment.create(env);

String db = "<DB Name>";

String warehouse = "<Warehouse Path>";

Map<String, String> props = new HashMap<>();
props.put("type", "iceberg");
props.put("warehouse", warehouse);
props.put("io-impl", "org.apache.iceberg.aws.s3.S3FileIO");

CatalogLoader glueCatlogLoader =
        CatalogLoader.custom(
                "glue",
                props,
                new Configuration(),
                "org.apache.iceberg.aws.glue.GlueCatalog");

TableLoader tableLoader =
        TableLoader.fromCatalog(glueCatlogLoader, TableIdentifier.of(db, "sample"));
```

```
DataStream<RowData> batch =

 FlinkSource.forRowData().env(env).tableLoader(tableLoader).streaming(false).build();

batch.print().name("print-sink");
```

Utilisation du catalogue Hive

Assurez-vous que les dépendances de Flink et Hive sont résolues comme décrit dans Configuration de Flink avec Hive Metastore et Glue Catalog.

Exécution d'une tâche Flink

Une façon de soumettre une tâche à Flink consiste à utiliser une session Flink YARN par tâche. Vous pouvez le lancer à l'aide de la commande suivante :

```
sudo flink run -m yarn-cluster -p 4 -yjm 1024m -ytm 4096m $JAR_FILE_NAME
```

Considérations relatives à l'utilisation d'Iceberg avec Flink

- Lorsque vous utilisez AWS Glue comme catalogue pour Iceberg, assurez-vous que la base de données dans laquelle vous créez une table existe dans AWS Glue. Si vous utilisez des services tels que AWS Lake Formation et que vous ne parvenez pas à charger le catalogue, assurez-vous de disposer d'un accès approprié au service pour exécuter la commande.
- L'intégration d'Iceberg Glue ne fonctionne pas avec le catalogue Redshift Managed Storage.

Utiliser un cluster Iceberg avec Hive

Avec les versions 6.9.0 et ultérieures d'Amazon EMR, vous pouvez utiliser Iceberg avec un cluster Hive sans avoir à effectuer les étapes de configuration requises pour l'intégration Open Source d'Iceberg Hive. Pour les versions 6.8.0 et antérieures d'Amazon EMR, vous pouvez utiliser une action d'amorçage pour installer le JAR `iceberg-hive-runtime` afin de configurer le support de Hive for Iceberg.

Amazon EMR 6.9.0 inclut toutes les fonctionnalités nécessaires à l'intégration de Hive 3.1.3 à Iceberg 0.14.1 et inclut également des fonctionnalités supplémentaires d'Amazon EMR telles que la sélection automatique des moteurs d'exécution pris en charge lors de l'exécution (Amazon EMR sur EKS 6.9.0).

Création d'un cluster Iceberg

Vous pouvez créer un cluster avec Iceberg installé à l'aide de l' AWS Management Console API, AWS CLI ou de l'API Amazon EMR. Dans ce didacticiel, vous allez utiliser le AWS CLI pour travailler avec Iceberg sur un cluster Amazon EMR. Pour utiliser la console afin de créer un cluster avec Iceberg installé, suivez les étapes de la section Création d'un lac de données Iceberg à l'aide d'Amazon Athena, d'Amazon EMR et d' AWS Glue.

Pour utiliser Iceberg sur Amazon EMR avec AWS CLI le, créez d'abord un cluster en suivant les étapes ci-dessous. Pour plus d'informations sur la spécification de la classification Iceberg à l'aide du SDK Java AWS CLI ou du SDK Java, consultez Fournir une configuration à l'aide de l' AWS CLI option lorsque vous créez un cluster ou. Fournir une configuration à l'aide du kit SDK Java lors de la création d'un cluster Créez un fichier nommé `configurations.json` avec le contenu suivant:

```
[{
    "Classification":"iceberg-defaults",
    "Properties":{"iceberg.enabled":"true"}
}]
```

Créez ensuite un cluster avec la configuration suivante, en remplaçant l'exemple de chemin de compartiment Amazon S3 et l'ID de sous-réseau par les vôtres :

```
aws emr create-cluster --release-label emr-6.9.0 \
--applications Name=Hive \
--configurations file://iceberg_configurations.json \
--region us-east-1 \
--name My_hive_Iceberg_Cluster \
--log-uri s3://amzn-s3-demo-bucket/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef
```

Un cluster Hive Iceberg effectue les opérations suivantes :

- Charge le fichier JAR d'exécution Iceberg Hive dans Hive et active la configuration associée à Iceberg pour le moteur Hive.
- Active la sélection dynamique du moteur d'exécution d'Amazon EMR Hive afin d'empêcher les utilisateurs de configurer le moteur d'exécution pris en charge pour la compatibilité avec Iceberg.

> **Note**
>
> Les clusters Hive Iceberg ne sont actuellement pas compatibles avec AWS Glue Data Catalog. Le catalogue Iceberg par défaut est `HiveCatalog`, ce qui correspond au métastore configuré pour l'environnement Hive. Pour plus d'informations sur la gestion des catalogues, consultez la section Utilisation HCatalog dans la documentation d'Apache Hive.

Prise en charge des fonctionnalités

Amazon EMR 6.9.0 prend en charge Hive 3.1.3 et Iceberg 0.14.1. La prise en charge des fonctionnalités est limitée aux fonctionnalités compatibles avec Iceberg pour Hive 3.1.2 et 3.1.3. Les commandes suivantes sont prises en charge :

- Avec les versions 6.9.0 à 6.12.x d'Amazon EMR, vous devez inclure le fichier jar `libfb303` dans le répertoire `auxlib` Hive. Utilisez la commande suivante pour l'inclure :

  ```
  sudo /usr/bin/ln -sf /usr/lib/hive/lib/libfb303-*.jar /usr/lib/hive/auxlib/
  libfb303.jar
  ```

  Avec les versions 6.13 et ultérieures d'Amazon EMR, le fichier jar `libfb303` est automatiquement lié symboliquement au répertoire `auxlib` Hive.

- Création d'une table
  - Table non partitionnée – Les tables externes dans Hive peuvent être créées en fournissant le gestionnaire de stockage comme suit :

    ```
    CREATE EXTERNAL TABLE x (i int) STORED BY
     'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
    ```

  - Table partitionnée – Les tables partitionnées externes dans Hive peuvent être créées comme suit :

    ```
    CREATE EXTERNAL TABLE x (i int) PARTITIONED BY (j int) STORED BY
     'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
    ```

> **Note**
>
> Le format de `STORED AS` fichier de n'ORC/AVRO/PARQUETest pas pris en charge dans Hive 3. La seule option par défaut est Parquet.

- Suppression d'une table – La commande `DROP TABLE` est utilisée pour supprimer des tables, comme dans l'exemple suivant :

```
DROP TABLE [IF EXISTS] table_name [PURGE];
```

- Lecture d'un tableau : Les instructions `SELECT` peuvent être utilisées pour lire les tables Iceberg dans Hive, comme dans l'exemple suivant. Les moteurs d'exécution pris en charge sont MR et Tez.

```
SELECT * FROM table_name
```

  Pour plus d'informations sur la syntaxe de sélection de Hive, voir LanguageManual Select. Pour plus d'informations sur les instructions select associées aux tables Iceberg dans Hive, consultez Apache Iceberg Select.

- Insertion dans une table – L'instruction de HiveQL `INSERT INTO` fonctionne sur les tables Iceberg avec le support du moteur d'exécution Map Reduce uniquement. Les utilisateurs d'Amazon EMR n'ont pas besoin de définir explicitement le moteur d'exécution, car Amazon EMR Hive sélectionne le moteur pour Iceberg Tables au moment de l'exécution.
  - Insert into d'une seul tableau – Exemple :

```
INSERT INTO table_name VALUES ('a', 1);
INSERT INTO table_name SELECT...;
```

  - Insert into multitables – Les instructions d'insertion multitables non atomiques dans des instructions sont prises en charge. Exemple :

```
FROM source
 INSERT INTO table_1 SELECT a, b
 INSERT INTO table_2 SELECT c,d;
```

À partir d'Amazon EMR 7.3.0, Hive with Iceberg prend en charge le catalogue de données AWS Glue en tant que métastore. Pour utiliser le catalogue de données AWS Glue comme métastore, définissez la propriété suivante.

```
SET iceberg.catalog.<catalog_name>.catalog-
impl=org.apache.iceberg.aws.glue.GlueCatalog;
```

Vous pouvez également définir la propriété suivante.

```
SET iceberg.catalog.<catalog_name>.type=glue;
```

Vous pouvez ensuite créer une table à l'aide de l'exemple suivant.

```
CREATE EXTERNAL TABLE table_name (col1 type1, col2 type2,..)
ROW FORMAT SERDE 'org.apache.iceberg.mr.hive.HiveIcebergSerDe'
STORED BY 'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
location '<location>'
TBLPROPERTIES ('table_type'='iceberg', 'iceberg.catalog'='<catalog_name>');
```

Considérations relatives à l'utilisation d'Iceberg avec Hive

- Iceberg prend en charge les types de requêtes suivants :
  - Create table
  - Drop table
  - Insérer dans la table
  - Lire la table

- Seul le moteur d'exécution MR (MapReduce) est pris en charge pour les opérations DML (langage de manipulation de données), et le MR est obsolète dans Hive 3.1.3.

- AWS Glue Data Catalog n'est actuellement pas compatible avec Iceberg with Hive.

- La gestion des erreurs n'est pas suffisamment robuste. En cas de mauvaise configuration, les insertions dans les requêtes peuvent se terminer correctement. Cependant, le fait de ne pas mettre à jour les métadonnées peut entraîner des pertes de données.
- L'intégration d'Iceberg Glue ne fonctionne pas avec le catalogue Redshift Managed Storage.

Historique des versions d'Iceberg

Le tableau suivant répertorie les versions d'Iceberg incluses dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur les versions d'Iceberg

Étiquette de version Amazon EMR	Version Iceberg	Composants installés avec Iceberg
emr-7.9.0	1.7.1-amzn-2	Not available.
emr-7,8.0	1.7.1-amzn-1	Not available.
emr-7,7.0	1.7.1-amzn-0	Not available.
emr-7.6.0	1.6.1-amzn-2	Not available.
emr-7.5.0	1.6.1-amzn-1	Not available.
emr-7.4.0	1.6.1-amzn-0	Not available.
emr-7.3.0	1.5.2-amzn-0	Not available.
emr-7.2.0	1,5.0-amzn-0	Not available.
emr-7.1.0	1.4.3-amzn-0	Not available.
emr-7.0.0	1.4.2-amzn-0	Not available.
emr-6.15.0	1.4.0-amzn-0	Not available.
emr-6.14.0	1.3.1-amzn-0	Not available.
emr-6.13.0	1.3.0-amzn-1	Not available.
emr-6.12.0	1.3.0-amzn-0	Not available.
emr-6.11.1	1.2.0-amzn-0	Not available.

Étiquette de version Amazon EMR	Version Iceberg	Composants installés avec Iceberg
emr-6.11.0	1.2.0-amzn-0	Not available.
emr-6.10.1	1.1.0-amzn-0	Not available.
emr-6.10.0	1.1.0-amzn-0	Not available.
emr-6.9.1	0.14.1-amzn-0	Not available.
emr-6.9.0	0.14.1-amzn-0	Not available.
emr-6.8.1	0.14.0-amzn-0	Not available.
emr-6.8.0	0.14.0-amzn-0	Not available.
emr-6.7.0	0.13.1-amzn-0	Not available.
emr-6.6.0	0.13.1	Not available.
emr-6.5.0	0.12.0	Not available.

Notes de mise à jour d'Iceberg par version

- Amazon EMR 6.9.0 - Notes de mise à jour d'Iceberg

Amazon EMR 6.9.0 - Notes de mise à jour d'Iceberg

Amazon EMR 6.9.0 - Changements apportés par Iceberg

Type	Description
Fonctionnalité	Intégration d'Amazon EMR Flink à Iceberg.
Fonctionnalité	Intégration de Hive d'Amazon EMR avec Iceberg.

Type	Description
Fonctionnalité	Support pour mettre en cache les fichiers de métadonnées Iceberg sur Amazon FSx pour Lustre afin d'améliorer le temps de planification des requêtes.
Rétroportage	PR 5050 : Flink 1.15 : Support des options d'écriture dans les commentaires SQL d'inserti on en ligne.
Rétroportage	PR 5282 : AWS : Corrigez les échecs de nouvelle tentative PUT en ouvrant de nouveaux flux de fichiers de données.
Rétroportage	PR 5318 : Flink 1.15 : Combler le fossé entre FlinkSource et IcebergSource (FLIP-27) et ajout d'une configuration opt-in pour utiliser la source FLIP-27 dans Flink SQL.
Rétroportage	PR 5344 : Flink 1.14 : Combler le fossé entre FlinkSource et IcebergSource (FLIP-27) et ajout d'une configuration opt-in pour utiliser la source FLIP-27 dans Flink SQL.
Rétroportage	PR 5393 : Flink 1.14, 1.15 : évitez de convertir les métriques Iceberg en Flink dans le lecteur MetricContext de source FLIP-27.
Rétroportage	PR 5401 : Flink 1.14, 1.15 : IcebergSo urceReader groupe manquant dans PR #5393 pour les métriques du lecteur de source FLIP-27.
Rétroportage	PR 5679 : Spark 3.2, 3.3 : correction de la propagation de nullabilité pour le nœud. MergeRows

Type	Description
Rétroportage	PR 5860 : Spark 3.3 : Correctif QueryFailure lors de l'exécution RewriteManifestProcedure sur des tables partitionnées par Date.
Rétroportage	PR 5880 : Spark 3.3 : correction de la nullité dans les projections. merge-on-read
Rétroportage	PR 5917 : Spark 3.2 : correction de la nullité dans les projections. merge-on-read

Bloc-notes Jupyter sur Amazon EMR

Le bloc-notes Jupyter est une application Web open source que vous pouvez utiliser pour créer et partager des documents qui contiennent du code interactif, des équations, des visualisations, et des textes narratifs. Amazon EMR vous offre trois options pour travailler avec les blocs-notes Jupyter :

Rubriques

- EMR Studio
- Bloc-notes Amazon EMR basé sur Jupyter Notebook
- JupyterHub

EMR Studio

Amazon EMR Studio est un environnement de développement intégré (IDE) basé sur le Web pour les blocs-notes Jupyter entièrement gérés qui s'exécutent sur des clusters Amazon EMR. Vous pouvez configurer un studio EMR pour que votre équipe développe, visualise et débogue des applications écrites en R, Python, Scala et. PySpark

Nous vous recommandons d'utiliser EMR Studio lorsque vous utilisez des blocs-notes Jupyter sur Amazon EMR. Pour plus d'informations, consultez EMR Studio dans le Guide de gestion d'Amazon EMR.

Bloc-notes Amazon EMR basé sur Jupyter Notebook

Blocs-notes EMR est un environnement Bloc-notes Jupyter intégré à la console Amazon EMR qui vous permet de créer rapidement des blocs-notes Jupyter, de les associer à des clusters Spark, puis d'ouvrir l'éditeur Jupyter Notebook dans la console pour exécuter des requêtes et du code à distance. Un bloc-notes EMR est enregistré dans Amazon S3 indépendamment des clusters pour un stockage durable, un accès rapide et une flexibilité accrue. Vous pouvez ouvrir plusieurs blocs-notes, associer plusieurs blocs-notes à un seul cluster et réutiliser un bloc-notes sur différents clusters.

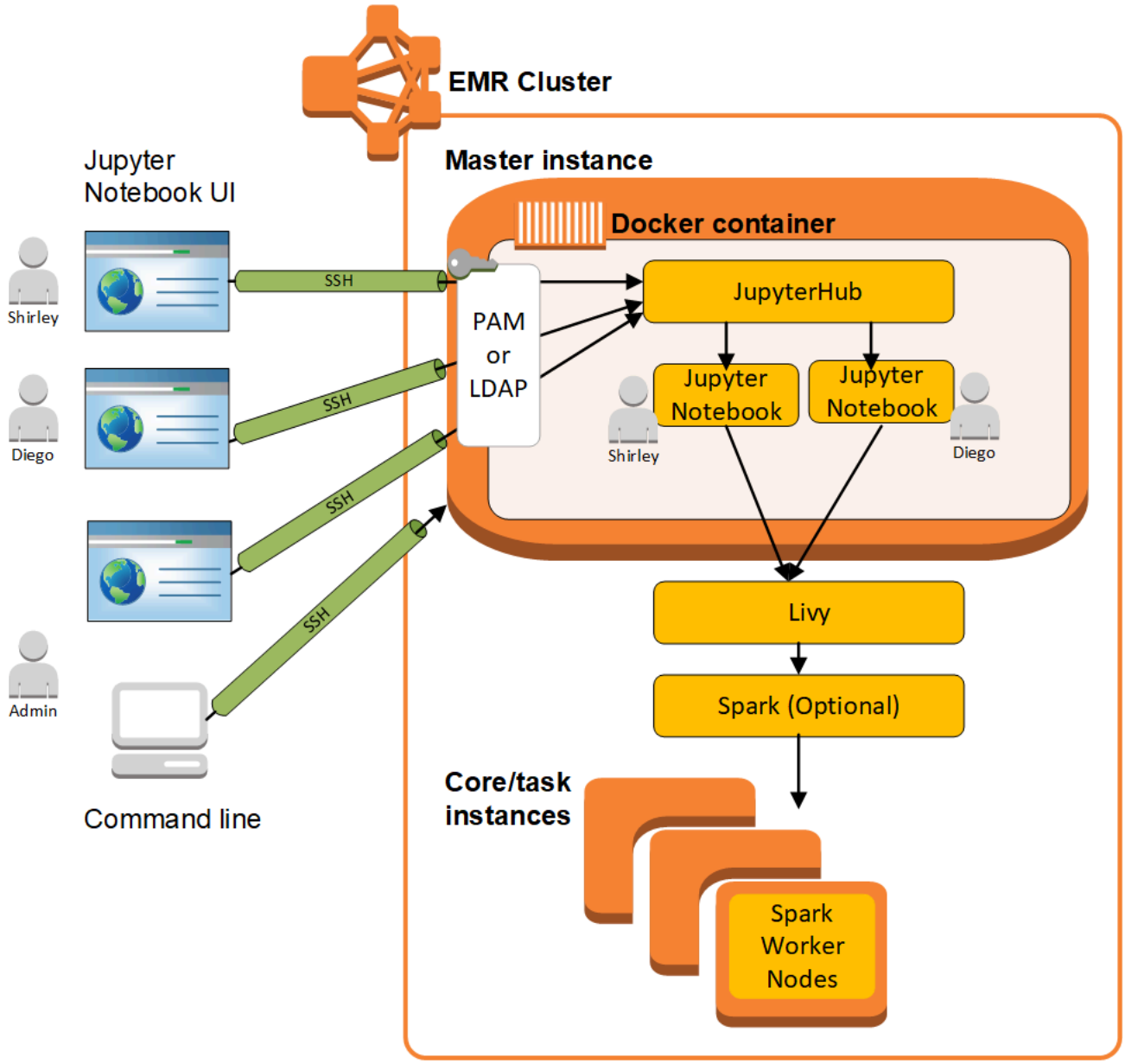
Pour plus d'informations, consultez Blocs-notes EMR dans le Guide de gestion d'Amazon EMR.

JupyterHub

Le bloc-notes Jupyter est une application Web open source que vous pouvez utiliser pour créer et partager des documents qui contiennent du code interactif, des équations, des visualisations, et des textes narratifs. JupyterHubvous permet d'héberger plusieurs instances d'un serveur de bloc-notes Jupyter pour utilisateur unique. Lorsque vous créez un cluster avec JupyterHub, Amazon EMR crée un conteneur Docker sur le nœud principal du cluster. JupyterHub, tous les composants requis pour Jupyter et Sparkmagic fonctionnent dans le conteneur.

Sparkmagic est une bibliothèque de noyaux qui permet aux blocs-notes Jupyter d'interagir avec Apache Spark exécuté sur Amazon EMR via Apache Livy, un serveur REST pour Spark. Spark et Apache Livy sont installés automatiquement lorsque vous créez un cluster avec JupyterHub. Le noyau Python 3 par défaut pour Jupyter est disponible en même temps que les noyaux PySpark 3 PySpark, et Spark, disponibles avec Sparkmagic. Vous pouvez utiliser ces noyaux pour exécuter le code Spark et les requêtes SQL interactives ad hoc à l'aide de Python et Scala. Vous pouvez installer manuellement des noyaux supplémentaires dans le conteneur Docker. Pour de plus amples informations, veuillez consulter Installation de noyaux et de bibliothèques supplémentaires.

Le schéma suivant illustre les composants de JupyterHub Amazon EMR avec des méthodes d'authentification correspondantes pour les utilisateurs de bloc-notes et l'administrateur. Pour de plus amples informations, veuillez consulter Ajout d'utilisateurs et d'administrateurs de bloc-notes Jupyter.



Le tableau suivant répertorie la version JupyterHub incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants avec lesquels Amazon EMR installe. JupyterHub

Pour connaître la version des composants installés avec JupyterHub cette version, consultez les versions des composants de la version 7.9.0.

JupyterHub informations de version pour emr-7.9.0

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-7.9.0	JupyterHub 1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Le tableau suivant répertorie la version JupyterHub incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants avec lesquels Amazon EMR installe. JupyterHub

Pour connaître la version des composants installés avec JupyterHub cette version, consultez les versions des composants de la version 6.15.0.

JupyterHub informations sur la version pour emr-6.15.0

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.15.0	JupyterHub 1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
		, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Le tableau suivant répertorie la version JupyterHub incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants avec lesquels Amazon EMR installe. JupyterHub

Pour connaître la version des composants installés avec JupyterHub cette version, consultez les versions des composants de la version 5.36.2.

JupyterHub informations de version pour emr-5.36.2

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.36.2	JupyterHub 1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Le noyau Python 3 inclus avec JupyterHub Amazon EMR est 3.6.4.

Les bibliothèques installées dans le `jupyterhub` conteneur peuvent varier selon les versions Amazon EMR et les versions Amazon EC2 AMI.

Pour lister les bibliothèques installées à l'aide de **`conda`**

- Exécutez la commande suivante sur la ligne de commande du nœud principal :

```
sudo docker exec jupyterhub bash -c "conda list"
```

Pour lister les bibliothèques installées à l'aide de **`pip`**

- Exécutez la commande suivante sur la ligne de commande du nœud principal :

```
sudo docker exec jupyterhub bash -c "pip freeze"
```

Rubriques

- Création d'un cluster avec JupyterHub
- Éléments à prendre en compte lors de l'utilisation JupyterHub sur Amazon EMR
- Configuration JupyterHub
- Configuration de la persistance pour les blocs-notes dans Amazon S3
- Connexion au nœud principal et aux serveurs de bloc-notes
- JupyterHub configuration et administration
- Ajout d'utilisateurs et d'administrateurs de bloc-notes Jupyter
- Installation de noyaux et de bibliothèques supplémentaires
- JupyterHub historique des publications

Création d'un cluster avec JupyterHub

Vous pouvez créer un cluster Amazon EMR à JupyterHub l'aide de la AWS Management Console AWS Command Line Interface, de l'ou de l'API Amazon EMR. Assurez-vous que le cluster n'est pas créé avec l'option d'arrêt automatique après exécution des étapes (option `--auto-terminate` dans l' AWS CLI). Assurez-vous également que les administrateurs et les utilisateurs du bloc-notes peuvent accéder à la paire de clés utilisée lors de la création du cluster. Pour plus d'informations, consultez Utilisation d'une paire de clés pour les informations d'identification SSH dans le Guide de gestion d'Amazon EMR.

Création d'un cluster à JupyterHub l'aide de la console

Utilisez la procédure suivante pour créer un cluster JupyterHub installé à l'aide de Options avancées dans la console Amazon EMR.

Pour créer un cluster Amazon EMR JupyterHub installé à l'aide de la console Amazon EMR

1. Accédez à la nouvelle console Amazon EMR et sélectionnez Changer pour l'ancienne console depuis le menu latéral. Pour plus d'informations sur ce qu'implique le passage à l'ancienne console, consultez la rubrique Utilisation de l'ancienne console.
2. Choisissez Créer un cluster et Go to advanced options (Aller aux options avancées).
3. Sous Software Configuration (Configuration logicielle) :
   - Pour Release, sélectionnez emr-5.36.2, puis choisissez. JupyterHub
   - Si vous utilisez Spark, pour utiliser le catalogue de données AWS Glue comme métastore pour Spark SQL, sélectionnez Utiliser pour les métadonnées des tables Spark. Pour de plus amples informations, veuillez consulter Utiliser le catalogue AWS Glue Data Catalog avec Spark sur Amazon EMR.
   - Pour Edit software settings (Modifier les paramètres logiciels), choisissez Enter configuration (Saisir une configuration) et spécifiez des valeurs, ou choisissez Load JSON from S3 (Charger JSON à partir de S3) et spécifiez un fichier de configuration JSON. Pour de plus amples informations, veuillez consulter Configuration JupyterHub.
4. Sous Add steps (optional) (Ajouter des étapes (facultatif)), configurez les étapes à exécuter lorsque le cluster est créé, assurez-vous que Auto-terminate cluster after the last step is completed (Arrêter automatiquement le cluster après l'exécution de la dernière étape) n'est pas sélectionné, puis choisissez Next (Suivant).
5. Choisissez les options Hardware Configuration (Configuration matérielle) et Next (Suivant). Pour plus d'informations, consultez Configuration du matériel et de la mise en réseau d'un cluster dans le Guide de gestion d'Amazon EMR.
6. Choisissez les options pour General Cluster Settings (Paramètres de cluster généraux) et Next (Suivant).
7. Choisissez Security Options (Options de sécurité) en spécifiant une paire de clés, puis choisissez Create Cluster (Créer un cluster).

Créez un cluster à JupyterHub l'aide du AWS CLI

Pour lancer un cluster avec JupyterHub, utilisez la `aws emr create-cluster` commande et, pour l'`--applications`option, spécifiez`Name=JupyterHub`. L'exemple suivant lance un JupyterHub cluster sur Amazon EMR avec deux EC2 instances (une instance principale et une instance de base). En outre, le débogage est activé et les journaux sont stockés dans l'emplacement Amazon S3, comme spécifié par `--log-uri`. La paire de clés spécifiée fournit l'accès aux EC2 instances Amazon dans le cluster.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
--applications Name=JupyterHub --log-uri s3://amzn-s3-demo-bucket/MyJupyterClusterLogs
 \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
 KeyName=MyKeyPair
```

Éléments à prendre en compte lors de l'utilisation JupyterHub sur Amazon EMR

Tenez compte des éléments suivants lorsque vous utilisez JupyterHub Amazon EMR.

- > **Warning**
  >
  > Les blocs-notes utilisateur et les fichiers sont enregistrées dans le système de fichiers sur le nœud principal. Il s'agit d'un stockage éphémère qui n'est pas conservé en cas d'arrêt du cluster. Lors de l'arrêt du cluster, ces données sont perdues si elles ne sont pas sauvegardées. Nous vous recommandons de planifier des sauvegardes régulières en utilisant des tâches `cron` ou tout autre méthode adaptée à votre application.
  > De plus, les changements de configuration effectués dans le conteneur peuvent ne pas être conservés lors du redémarrage de celui-ci. Nous vous recommandons de créer des

scripts ou d'automatiser de quelque manière que ce soit les configurations de conteneur, de sorte que vous puissiez reproduire les personnalisations plus rapidement.

- L'authentification Kerberos configurée à l'aide d'une configuration de sécurité Amazon EMR n'est pas prise en charge.
- OAuthenticator n'est pas pris en charge.

Configuration JupyterHub

Vous pouvez personnaliser la configuration JupyterHub sur Amazon EMR et les blocs-notes pour utilisateur individuel en vous connectant au nœud principal du cluster et en modification les fichiers de configuration. Après avoir modifié les valeurs, redémarrez le conteneur `jupyterhub`.

Modifiez les propriétés dans les fichiers suivants pour configurer JupyterHub et personnaliser les blocs-notes Jupyter :

- `jupyterhub_config.py` – Par défaut, ce fichier est enregistré dans le répertoire `/etc/jupyter/conf/` sur le nœud principal. Pour plus d'informations, consultez Principes de base de configuration dans la JupyterHub documentation.
- `jupyter_notebook_config.py` – Ce fichier est enregistré dans le répertoire `/etc/jupyter/` par défaut et copié dans le conteneur `jupyterhub` en tant que valeur par défaut. Pour plus d'informations, consultez Fichier de configuration et options de ligne de commande dans la documentation sur les bloc-notes Jupyter.

Vous pouvez également utiliser la classification de configuration `jupyter-sparkmagic-conf` pour personnaliser Sparkmagic, ce qui met à jour les valeurs du fichier `config.json` pour Sparkmagic. Pour plus d'informations sur les paramètres disponibles, consultez le fichier example_config.json sur. GitHub Pour plus d'informations sur l'utilisation des classifications de configuration avec des applications dans Amazon EMR, consultez Configuration des applications.

L'exemple suivant lance un cluster en utilisant le AWS CLI, en référençant le fichier `MyJupyterConfig.json` pour les paramètres de classification de configuration de Sparkmagic.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --use-default-roles --release-label emr-5.14.0 \
--applications Name=Jupyter --instance-type m4.xlarge --instance-count 3 \
--ec2-attributes KeyName=MyKey,SubnetId=subnet-1234a5b6 --configurations file://
MyJupyterConfig.json
```

Exemples de contenu du fichier `MyJupyterConfig.json` :

```
[
    {
    "Classification":"jupyter-sparkmagic-conf",
    "Properties": {
      "kernel_python_credentials" : "{\"username\":\"diego\",\"base64_password\":
\"mypass\",\"url\":\"http:\/\/localhost:8998\",\"auth\":\"None\"}"
      }
    }
]
```

> **Note**
>
> Avec la version 5.21.0 et ultérieures d'Amazon EMR, vous permet de remplacer les configurations de cluster et de spécifier des classifications de configuration supplémentaires pour chaque groupe d'instances dans un cluster en cours d'exécution. Pour ce faire, utilisez la console Amazon EMR, le AWS Command Line Interface (AWS CLI) ou le AWS SDK. Pour plus d'informations, consultez Fourniture d'une configuration pour un groupe d'instances dans un cluster en cours d'exécution.

Configuration de la persistance pour les blocs-notes dans Amazon S3

Vous pouvez configurer un JupyterHub cluster dans Amazon EMR afin que les blocs-notes enregistrés par un utilisateur persistent dans Amazon S3, en dehors de stockage éphémère sur les instances de cluster. EC2

Vous spécifiez une persistence Amazon S3 à l'aide de la classification de configuration `jupyter-s3-conf` lorsque vous créez un cluster. Pour de plus amples informations, veuillez consulter Configuration des applications.

En plus de l'activation de persistance Amazon S3 à l'aide de la propriété `s3.persistence.enabled`, vous devez spécifier un compartiment dans Amazon S3 où les blocs-notes sont enregistrés à l'aide de la propriété `s3.persistence.bucket`. Les blocs-notes de chaque utilisateur sont enregistrés dans un dossier `jupyter/`*`jupyterhub-user-name`* dans le compartiment spécifié. Le compartiment doit déjà exister dans Amazon S3, et le rôle pour le profil d' EC2 instance que vous spécifiez lorsque vous créez le cluster doit disposer des autorisations pour le compartiment (par défaut, le rôle est`EMR_EC2_DefaultRole`). Pour plus d'informations, consultez Configuration de rôles IAM pour des autorisations AWS Amazon EMR aux services.

Lorsque vous lancez un nouveau cluster à l'aide des mêmes propriétés de classification de configuration, les utilisateurs peuvent ouvrir des blocs-notes avec le contenu à partir de l'emplacement enregistré.

Notez que lorsque vous importez des fichiers sous forme de modules dans un bloc-notes lorsque Amazon S3 est activé, cela entraîne le chargement des fichiers sur Amazon S3. Lorsque vous importez des fichiers sans activer la persistance Amazon S3, ils sont chargés JupyterHub dans votre conteneur.

L'exemple suivant active la persistance Amazon S3. Les cahiers enregistrés par les utilisateurs sont enregistrés dans le dossier `s3://MyJupyterBackups/jupyter/`*`jupyterhub-user-name`* pour chaque utilisateur, où *`jupyterhub-user-name`* est un nom d'utilisateur, par exemple `diego`.

```
[
    {
        "Classification": "jupyter-s3-conf",
        "Properties": {
            "s3.persistence.enabled": "true",
            "s3.persistence.bucket": "MyJupyterBackups"
        }
    }
]
```

Connexion au nœud principal et aux serveurs de bloc-notes

JupyterHub les administrateurs et les utilisateurs d'ordinateurs portables doivent se connecter au nœud principal du cluster à l'aide d'un tunnel SSH, puis se connecter aux interfaces Web desservies

par JupyterHub le nœud principal. Pour plus d'informations sur la configuration d'un tunnel SSH et l'utilisation de ce tunnel pour les connexions Web via un proxy, consultez Connexion au cluster dans le Guide de gestion d'Amazon EMR.

Par défaut, JupyterHub sur Amazon EMR est disponible via le port 9443 sur le nœud principal. Le JupyterHub proxy interne dessert également les instances de bloc-notes via le port 9443. JupyterHub et les interfaces Web de Jupyter sont accessibles à l'aide d'une URL suivant le modèle suivant :

https : //:9443 ***MasterNodeDNS***

Vous pouvez spécifier un autre port à l'aide de la propriété `c.JupyterHub.port` dans le fichier `jupyterhub_config.py`. Pour plus d'informations, consultez Principes de base de la mise en réseau dans la JupyterHub documentation.

Par défaut, JupyterHub sur Amazon EMR utilise un certificat auto-signé pour le chiffrement SSL avec le protocole HTTPS. Les utilisateurs sont invités à approuver le certificat auto-signé lorsqu'ils se connectent. Vous pouvez utiliser un certificat approuvé et des clés qui vous sont propres. Remplacez le fichier de certificat par défaut, `server.crt` et le fichier de clé `server.key` dans le répertoire `/etc/jupyter/conf/` du nœud principal par un certificat et des fichiers de clé qui vous sont propres. Utilisez les propriétés `c.JupyterHub.ssl_key` et `c.JupyterHub.ssl_cert` dans le fichier `jupyterhub_config.py` pour spécifier vos matériaux SSL. Pour plus d'informations, consultez la section Paramètres de sécurité dans la JupyterHub documentation. Après avoir mis à jour `jupyterhub_config.py`, redémarrez le conteneur.

JupyterHub configuration et administration

JupyterHub et les composants associés s'exécutent dans un conteneur Docker nommé `jupyterhub` qui exécute le système d'exploitation Ubuntu. Il existe plusieurs méthodes pour administrer des composants s'exécutant à l'intérieur du conteneur.

> **Warning**
>
> Les personnalisations que vous effectuez dans le conteneur peuvent ne pas être conservées si le conteneur redémarre. Nous vous recommandons de créer des scripts ou d'automatiser de quelque manière que ce soit les configurations de conteneur, de sorte que vous puissiez reproduire les personnalisations plus rapidement.

Administration à l'aide de la ligne de commande

Une fois connecté au nœud principal à l'aide de SSH, vous pouvez émettre des commandes en utilisant l'interface de ligne de commande Docker et en spécifiant le conteneur par l'intermédiaire de son nom (`jupyterhub`) ou de son ID. Par exemple, `sudo docker exec jupyterhub` *`command`* exécute des commandes reconnues par le système d'exploitation ou par une application exécutée dans le conteneur. Vous pouvez utiliser cette méthode pour ajouter des utilisateurs au système d'exploitation et installer des applications et des bibliothèques supplémentaires au sein du conteneur Docker. Par exemple, l'image de conteneur par défaut inclut Conda pour l'installation de package ; vous pouvez donc exécuter la commande suivante sur la ligne de commande du nœud principal pour installer une application, Keras, dans le conteneur :

```
sudo docker exec jupyterhub conda install keras
```

Administration par soumission d'étapes

Les étapes constituent un moyen de soumettre le travail à un cluster. Vous pouvez soumettre des étapes lors du lancement d'un cluster ou lorsque celui-ci est en cours d'exécution. Les commandes exécutées sur la ligne de commande peuvent être soumises sous forme d'étapes à l'aide de `command-runner.jar`. Pour plus d'informations, consultez Utilisation des étapes à l'aide de la CLI et de la console dans le Guide de gestion d'Amazon EMR et Exécuter des commandes et des scripts sur un cluster Amazon EMR.

Par exemple, vous pouvez utiliser la AWS CLI commande suivante sur un ordinateur local pour installer Keras de la même manière que vous l'avez fait à partir de la ligne de commande du nœud principal dans l'exemple précédent :

```
aws emr add-steps --cluster-id MyClusterID --steps Name="Command
 Runner",Jar="command-runner.jar",Args="/usr/bin/sudo","/usr/bin/
docker","exec","jupyterhub","conda","install","keras"
```

De plus, vous pouvez créer un script avec une séquence d'étapes, le charger sur Amazon S3, puis utiliser `script-runner.jar` pour l'exécuter lors de la création du cluster ou l'ajouter en tant qu'étape. Pour de plus amples informations, veuillez consulter Exécuter des commandes et des scripts sur un cluster Amazon EMR. Pour obtenir un exemple, consultez the section called “Exemple : script bash pour ajouter plusieurs utilisateurs”.

Administration à l'aide de REST APIs

Jupyter et JupyterHub le proxy HTTP JupyterHub fournissent un REST APIs que vous pouvez utiliser pour envoyer des demandes. Pour envoyer des demandes à JupyterHub, vous devez transmettre un jeton d'API avec la demande. Vous pouvez utiliser la commande `curl` à partir de l'interface de ligne de commande du nœud principal pour exécuter des commandes REST. Pour plus d'informations, consultez les ressources suivantes :

- Utilisation JupyterHub de l'API REST dans la documentation de JupyterHub, qui inclut des instructions pour générer des jetons d'API
- API de serveur de blocs-notes Jupyter sur GitHub
- configurable-http-proxysur GitHub

L'exemple suivant montre comment utiliser l'API REST JupyterHub pour obtenir une liste d'utilisateurs. La commande transmet un jeton d'administration généré précédemment et utilise le port par défaut, 9443 JupyterHub, pour rediriger la sortie vers jq afin de faciliter la visualisation :

```
curl -XGET -s -k https://$HOST:9443/hub/api/users \
-H "Authorization: token $admin_token" | jq .
```

Ajout d'utilisateurs et d'administrateurs de bloc-notes Jupyter

Vous pouvez utiliser l'une des deux méthodes permettant aux utilisateurs de s'authentifier JupyterHub afin qu'ils puissent créer des blocs-notes et, éventuellement, les administrer. JupyterHub La méthode la plus simple consiste à utiliser JupyterHub le module d'authentification enfichable (PAM). En outre, JupyterHub sur Amazon EMR prend en charge le Plug-in d'authentification LDAP JupyterHub pour l'obtention de l'identité des utilisateurs à partir d'un serveur LDAP, tel qu'un serveur Microsoft Active Directory. Des instructions et des exemples pour l'ajout d'utilisateurs avec chaque méthode d'authentification sont fournis dans cette section.

JupyterHub sur Amazon EMR a un utilisateur par défaut qui dispose des autorisations d'administrateur. Le nom d'utilisateur est `jovyan`, et le mot de passe est `jupyter`. Nous vous recommandons vivement de remplacer cet utilisateur par un autre utilisateur disposant des autorisations d'administrateur. Vous pouvez effectuer cette opération via une étape lorsque vous créez le cluster, ou en vous connectant au nœud maître lorsque le cluster est en cours d'exécution.

Rubriques

- Utilisation de l'authentification PAM
- Utilisation de l'authentification LDAP
- Emprunt d'identité de l'utilisateur

Utilisation de l'authentification PAM

La création d'utilisateurs PAM JupyterHub sur Amazon EMR est un processus en deux étapes. La première étape consiste à ajouter des utilisateurs au système d'exploitation qui s'exécute dans le conteneur `jupyterhub` sur le nœud principal, et d'ajouter le répertoire de base correspondant pour chaque utilisateur. La deuxième étape consiste à ajouter ces utilisateurs du système d'exploitation en tant qu' JupyterHubutilisateurs, un processus connu sous le nom de mise sur liste blanche dans. JupyterHub Une fois qu'un JupyterHub utilisateur est ajouté, il peut se connecter à l' JupyterHub URL et fournir les informations d'identification de son système d'exploitation pour y accéder.

Lorsqu'un utilisateur se connecte, JupyterHub ouvre l'instance de serveur de bloc-notes pour cet utilisateur, qui est enregistrée dans le répertoire personnel de l'utilisateur sur le nœud principal, à savoir`/var/lib/jupyter/home/`*`username`*. Si une instance de serveur de bloc-notes n'existe pas, JupyterHub génère une instance de bloc-notes dans le répertoire de base de l'utilisateur. Les sections suivantes montrent comment ajouter des utilisateurs individuellement au système d'exploitation et à JupyterHub, puis comment utiliser un script bash rudimentaire qui ajoute plusieurs utilisateurs.

Ajout d'un utilisateur du système d'exploitation au conteneur

L'exemple suivant utilise tout d'abord la commande useradd dans le conteneur pour ajouter un utilisateur unique, diego, et créer un répertoire de base pour cet utilisateur. La deuxième commande utilise chpasswd pour établir le mot de passe associé à diego pour cet utilisateur. Les commandes sont exécutées sur la ligne de commande du nœud principal lorsque vous êtes connecté à l'aide de SSH. Vous pouvez également exécuter ces commandes à l'aide d'une étape, comme décrit précédemment dans Administration par soumission d'étapes.

```
sudo docker exec jupyterhub useradd -m -s /bin/bash -N diego
sudo docker exec jupyterhub bash -c "echo diego:diego | chpasswd"
```

Ajouter un JupyterHub utilisateur

Vous pouvez utiliser le panneau d'administration JupyterHub ou l'API REST pour ajouter des utilisateurs et des administrateurs, ou simplement des utilisateurs.

Pour ajouter des utilisateurs et des administrateurs à l'aide du volet admin dans JupyterHub

1. Connectez-vous au nœud principal via SSH et connectez-vous à https : //:9443 *MasterNodeDNS* avec une identité dotée d'autorisations d'administrateur.
2. Choisissez Control Panel (Panneau de configuration), Admin.
3. Choisissez User (Utilisateur), Add Users (Ajouter des utilisateurs), ou choisissez Admin, Add Admins (Ajouter des administrateurs).

Pour ajouter un utilisateur à l'aide de l'API REST

1. Connectez-vous au nœud principal à l'aide de SSH et utilisez la commande suivante sur le nœud principal, ou exécutez la commande en tant qu'étape.
2. Procurez-vous un jeton administratif pour effectuer des demandes d'API et remplacez-le *AdminToken* à l'étape suivante par ce jeton.
3. Utilisez la commande suivante en *UserName* remplaçant par un utilisateur du système d'exploitation créé dans le conteneur.

```
curl -XPOST -H "Authorization: token AdminToken" "https://$(hostname):9443/hub/api/
users/UserName
```

> **Note**
>
> Vous êtes automatiquement ajouté en tant qu'utilisateur JupyterHub non administrateur lorsque vous vous connectez à l'interface JupyterHub Web pour la première fois.

Exemple : script bash pour ajouter plusieurs utilisateurs

L'exemple de script bash suivant associe les étapes précédentes de cette section pour créer plusieurs JupyterHub utilisateurs. Le script peut être exécuté directement sur le nœud principal, ou chargé sur Amazon S3 et exécuté en tant qu'étape.

Le script établit tout d'abord une liste des noms d'utilisateur, puis utilise la commande `jupyterhub token` pour créer un jeton d'API pour l'administrateur par défaut jovyan. Ensuite, il crée un utilisateur du système d'exploitation dans le conteneur `jupyterhub` pour chaque utilisateur, et lui affecte un mot de passe initial équivalent à son nom d'utilisateur. Enfin, il appelle l'opération de l'API REST pour

créer chaque utilisateur dans JupyterHub. Il transmet le jeton généré précédemment dans le script et transfère la réponse REST à `jq` pour faciliter l'affichage.

```
# Bulk add users to container and JupyterHub with temp password of username
set -x
USERS=(shirley diego ana richard li john mary anaya)
TOKEN=$(sudo docker exec jupyterhub /opt/conda/bin/jupyterhub token jovyan | tail -1)
for i in "${USERS[@]}";
do
   sudo docker exec jupyterhub useradd -m -s /bin/bash -N $i
   sudo docker exec jupyterhub bash -c "echo $i:$i | chpasswd"
   curl -XPOST --silent -k https://$(hostname):9443/hub/api/users/$i \
 -H "Authorization: token $TOKEN" | jq
done
```

Enregistrez le script dans un emplacement Amazon S3 tel que `s3://amzn-s3-demo-bucket/createjupyterusers.sh`. Vous pouvez ensuite utiliser `script-runner.jar` pour l'exécuter en tant qu'étape.

Exemple : exécution du script lors de la création d'un cluster (AWS CLI)

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
--applications Name=JupyterHub --log-uri s3://amzn-s3-demo-bucket/MyJupyterClusterLogs
 \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
 KeyName=MyKeyPair \
--steps Type=CUSTOM_JAR,Name=CustomJAR,ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://amzn-
s3-demo-bucket/createjupyterusers.sh"]
```

Exécution du script sur un cluster existant (AWS CLI)

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr add-steps --cluster-id j-XXXXXXXX --steps Type=CUSTOM_JAR,\
Name=CustomJAR,ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://amzn-s3-demo-bucket/createjupyterusers.sh"]
```

Utilisation de l'authentification LDAP

LDAP (Lightweight Directory Access Protocol) est un protocole d'application pour l'interrogation et la modification d'objets correspondant à des ressources telles que des utilisateurs et des ordinateurs stockés dans un fournisseur de services de répertoire compatible avec LDAP tel qu'Active Directory, ou dans un serveur OpenLDAP. Vous pouvez utiliser le Plug-in d'authentification LDAP pour with JupyterHub sur JupyterHub Amazon EMR pour utiliser LDAP pour l'authentification d'utilisateur. Le plug-in gère les sessions de connexion pour les utilisateurs LDAP et fournit des informations utilisateur à Jupyter. Cela permet aux utilisateurs de se connecter à JupyterHub des blocs-notes en utilisant les informations d'identification de leur identité stockées sur un serveur compatible LDAP.

Les étapes de cette section vous guident dans les étapes suivantes pour configurer et activer LDAP à l'aide du Plug-in d'authentification LDAP pour. JupyterHub Pour effectuer ces étapes, vous devez être connecté à la ligne de commande du nœud principal. Pour de plus amples informations, veuillez consulter Connexion au nœud principal et aux serveurs de bloc-notes.

1. Créez un fichier de configuration LDAP avec les informations sur le serveur LDAP, telles que l'adresse IP de l'hôte, le port, les noms de liaison, etc.
2. Modifiez `/etc/jupyter/conf/jupyterhub_config.py` pour activer le plug-in d'authentification LDAP pour JupyterHub.
3. Créez et exécutez un script qui configure LDAP dans le conteneur `jupyterhub`.
4. Interrogez LDAP pour les utilisateurs, puis créez des répertoires personnels dans le conteneur pour chaque utilisateur. JupyterHub nécessite des répertoires personnels pour héberger les blocs-notes.

5. Exécuter un script qui redémarre JupyterHub

> **Important**
>
> Avant de configurer LDAP, testez votre infrastructure réseau afin de vous assurer que le serveur LDAP et le nœud principal du cluster peuvent communiquer si nécessaire. Généralement, TLS utilise le port 389 sur une connexion TCP brute. Si votre connexion LDAP utilise SSL, le port TCP connu pour SSL est 636.

Création du fichier de configuration LDAP

L'exemple ci-dessous utilise les valeurs de configuration d'espace réservé suivantes. Remplacez ces valeurs par des paramètres qui correspondent à votre implémentation.

- Le serveur LDAP exécute la version 3 et est disponible sur le port 389. Il s'agit du port non SSL standard pour LDAP.
- Le nom unique de base est `dc=example, dc=org`.

Utilisez un éditeur de texte pour créer le fichier ldap.conf, avec un contenu similaire à ce qui suit. Utilisez les valeurs appropriées pour votre implémentation LDAP. *`host`*Remplacez-le par l'adresse IP ou le nom d'hôte résoluble de votre serveur LDAP.

```
base dc=example,dc=org
uri ldap://host
ldap_version 3
binddn cn=admin,dc=example,dc=org
bindpw admin
```

Activation du plug-in d'authentification LDAP pour JupyterHub

Utilisez un éditeur de texte pour modifier le fichier `/etc/jupyter/conf/jupyterhub_config.py` et ajouter les propriétés ldapauthenticator similaires à ce qui suit. Remplacez *`host`* par l'adresse IP ou le nom d'hôte résoluble du serveur LDAP. L'exemple suppose que les objets utilisateur se trouvent au sein d'une unité organisationnelle (ou) nommée *`people`* et utilise les composants de nom unique que vous avez définis précédemment`ldap.conf`.

```
c.JupyterHub.authenticator_class = 'ldapauthenticator.LDAPAuthenticator'
```

```
c.LDAPAuthenticator.use_ssl = False
c.LDAPAuthenticator.server_address = 'host'
c.LDAPAuthenticator.bind_dn_template = 'cn={username},ou=people,dc=example,dc=org'
```

Configuration de LDAP dans le conteneur

Utilisez un éditeur de texte pour créer un script bash avec le contenu suivant :

```
#!/bin/bash

# Uncomment the following lines to install LDAP client libraries only if
# using Amazon EMR release version 5.14.0. Later versions install libraries by default.
# sudo docker exec jupyterhub bash -c "sudo apt-get update"
# sudo docker exec jupyterhub bash -c "sudo apt-get -y install libnss-ldap libpam-ldap
 ldap-utils nscd"

# Copy ldap.conf
sudo docker cp ldap.conf jupyterhub:/etc/ldap/
sudo docker exec jupyterhub bash -c "cat /etc/ldap/ldap.conf"

# configure nss switch
sudo docker exec jupyterhub bash -c "sed -i 's/\(^passwd.*\)/\1 ldap/g' /etc/
nsswitch.conf"
sudo docker exec jupyterhub bash -c "sed -i 's/\(^group.*\)/\1 ldap/g' /etc/
nsswitch.conf"
sudo docker exec jupyterhub bash -c "sed -i 's/\(^shadow.*\)/\1 ldap/g' /etc/
nsswitch.conf"
sudo docker exec jupyterhub bash -c "cat /etc/nsswitch.conf"

# configure PAM to create home directories
sudo docker exec jupyterhub bash -c "echo 'session required        pam_mkhomedir.so
 skel=/etc/skel umask=077' >> /etc/pam.d/common-session"
sudo docker exec jupyterhub bash -c "cat /etc/pam.d/common-session"

# restart nscd service
sudo docker exec jupyterhub bash -c "sudo service nscd restart"

# Test
sudo docker exec jupyterhub bash -c "getent passwd"

# Install ldap plugin
sudo docker exec jupyterhub bash -c "pip install jupyterhub-ldapauthenticator"
```

Enregistrez le script sur le nœud principal, puis exécutez-le à partir de la ligne de commande du nœud principal. Par exemple, avec le script enregistré sous le nom `configure_ldap_client.sh`, rendez le fichier exécutable :

```
chmod +x configure_ldap_client.sh
```

Exécutez ensuite le script :

```
./configure_ldap_client.sh
```

Ajout d'attributs à Active Directory

Pour rechercher chaque utilisateur et créer l'entrée appropriée dans la base de données, le conteneur JupyterHub docker nécessite les propriétés UNIX suivantes pour l'objet utilisateur correspondant dans Active Directory. Pour plus d'informations, consultez Comment puis-je continuer à modifier les attributs GID/UID RFC 2307 maintenant que le plug-in Attributs Unix n'est plus disponible pour le composant logiciel enfichable Utilisateurs Active Directory et Ordinateurs MMC ? dans l'article Éclaircissement concernant le statut de la gestion des identités pour Unix (IDMU) et le rôle de serveur NIS dans l'aperçu technique de Windows Server 2016 et au-delà.

- `homeDirectory`

  Il s'agit de l'emplacement de l'annuaire de base de l'utilisateur, qui est généralement `/home/`*`username`*.

- `gidNumber`

  Il s'agit d'une valeur supérieure à 60 000 qui n'est pas déjà utilisée par un autre utilisateur. Vérifiez le fichier `etc/passwd` pour les GID en cours d'utilisation.

- `uidNumber`

  Il s'agit d'une valeur supérieure à 60 000 qui n'est pas déjà utilisée par un autre groupe. Vérifiez le fichier `etc/group` pour les UID en cours d'utilisation.

- `uid`

  C'est le même que le*`username`*.

Création de répertoires de base utilisateur

JupyterHub a besoin de répertoires de base dans le conteneur pour authentifier les utilisateurs LDAP et stocker les données d'instance. L'exemple suivant présente deux utilisateurs, shirley et diego, dans le répertoire LDAP.

La première étape consiste à interroger le serveur LDAP pour obtenir les informations d'identifiant d'utilisateur et d'identifiant de groupe de chaque utilisateur à l'aide de ldapsearch, comme indiqué dans l'exemple suivant, en les *host* remplaçant par l'adresse IP ou le nom d'hôte résoluble de votre serveur LDAP :

```
ldapsearch -x -H ldap://host \
 -D "cn=admin,dc=example,dc=org" \
 -w admin \
 -b "ou=people,dc=example,dc=org" \
 -s sub \
 "(objectclass=*)" uidNumber gidNumber
```

La commande `ldapsearch` renvoie une réponse au format LDIF qui ressemble à ce qui suit pour les utilisateurs shirley et diego.

```
# extended LDIF

# LDAPv3
# base <ou=people,dc=example,dc=org> with scope subtree
# filter: (objectclass=*)
# requesting: uidNumber gidNumber sn

# people, example.org
dn: ou=people,dc=example,dc=org

# diego, people, example.org
dn: cn=diego,ou=people,dc=example,dc=org
sn: B
uidNumber: 1001
gidNumber: 100

# shirley, people, example.org
dn: cn=shirley,ou=people,dc=example,dc=org
sn: A
uidNumber: 1002
gidNumber: 100
```

```
# search result
search: 2
result: 0 Success

# numResponses: 4
# numEntries: 3
```

À l'aide des informations contenues dans la réponse, exécutez des commandes dans le conteneur de manière à créer un répertoire de base pour chaque nom commun d'utilisateur (`cn`). Utilisez la valeur de `uidNumber` et de `gidNumber` pour corriger la propriété du répertoire de base pour cet utilisateur. Les exemples de commandes suivants le font pour l'utilisateur*shirley*.

```
sudo docker container exec jupyterhub bash -c "mkdir /home/shirley"
sudo docker container exec jupyterhub bash -c "chown -R $uidNumber /home/shirley"
sudo docker container exec jupyterhub bash -c "sudo chgrp -R $gidNumber /home/shirley"
```

> **Note**
>
> L'authentificateur LDAP pour JupyterHub ne prend pas en charge la création d'utilisateurs locaux. Pour plus d'informations, consultez la Remarque sur la configuration de l'authentificateur LDAP pour la création d'utilisateurs locaux.
> Pour créer manuellement un utilisateur local, utilisez la commande suivante.
>
> ```
> sudo docker exec jupyterhub bash -c "echo 'shirley:x:$uidNumber:$gidNumber::/
> home/shirley:/bin/bash' >> /etc/passwd"
> ```

Redémarrage du JupyterHub conteneur

Exécutez la commande suivante pour redémarrer le conteneur `jupyterhub` :

```
sudo docker stop jupyterhub
sudo docker start jupyterhub
```

Emprunt d'identité de l'utilisateur

Une tâche Spark exécutée dans un bloc-note Jupyter traverse plusieurs applications au cours de son exécution sur Amazon EMR. Par exemple, PySpark 3 codes exécutés par un utilisateur dans Jupyter

sont reçus par Sparkmagic, qui utilise une requête HTTP POST pour le soumettre à Livy, qui crée ensuite une tâche Spark à exécuter sur le cluster à l'aide de YARN.

Par défaut, les tâches YARN soumises de cette manière s'exécutent sous l'utilisateur `livy`, quel que soit l'utilisateur ayant lancé la tâche. En configurant l'emprunt d'identité de l'utilisateur, l'ID utilisateur de l'utilisateur du bloc-notes peut également être l'utilisateur associé à la tâche YARN. Plutôt que les tâches soient lancées par les deux utilisateurs `shirley` et `diego` associés à l'utilisateur `livy`, les tâches initiées par chaque utilisateur sont associées respectivement à `shirley` et à `diego`. Cela vous permet de vérifier l'utilisation de Jupyter et de gérer les applications au sein de votre organisation.

Cette configuration est prise en charge uniquement lorsque des appels de Sparkmagic vers Livy ne sont pas authentifiés. Les applications qui fournissent une couche d'authentification ou de redirection entre les applications Hadoop et Livy (par exemple, Apache Knox Gateway) ne sont pas prises en charge. Les étapes de configuration de l'emprunt d'identité utilisateur décrites dans cette section supposent que Livy JupyterHub et Livy s'exécutent sur le même nœud principal. Si votre application a des clusters séparés, Étape 3 : Créer des répertoires de base HDFS pour les utilisateurs doit être modifiée afin que les annuaires HDFS soient créés sur le nœud maître Livy.

Étapes pour configurer l'emprunt d'identité de l'utilisateur

- Étape 1 : Configurer Livy
- Étape 2 : Ajouter des utilisateurs
- Étape 3 : Créer des répertoires de base HDFS pour les utilisateurs

Étape 1 : Configurer Livy

Vous utilisez les classifications de configuration `livy-conf` et `core-site` lorsque vous créez un cluster pour activer l'emprunt d'identité de l'utilisateur Livy, comme illustré dans l'exemple suivant. Enregistrez la classification de configuration sous la forme d'un fichier JSON, puis faites-y référence lorsque vous créez le cluster ou que vous spécifiez la classification de configuration en ligne. Pour de plus amples informations, veuillez consulter Configuration des applications.

```
[
  {
    "Classification": "livy-conf",
    "Properties": {
      "livy.impersonation.enabled": "true"
    }
```

```
  },
  {
    "Classification": "core-site",
    "Properties": {
      "hadoop.proxyuser.livy.groups": "*",
      "hadoop.proxyuser.livy.hosts": "*"
    }
  }
]
```

Étape 2 : Ajouter des utilisateurs

Ajoutez des JupyterHub utilisateurs à l'aide de PAM ou LDAP. Pour plus d'informations, consultez Utilisation de l'authentification PAM et Utilisation de l'authentification LDAP.

Étape 3 : Créer des répertoires de base HDFS pour les utilisateurs

Vous vous êtes connecté au nœud maître pour créer des utilisateurs. Toujours connecté au nœud maître, copiez le contenu ci-dessous et enregistrez-le dans un fichier script. Le script crée des répertoires de base HDFS pour chaque JupyterHub utilisateur du nœud principal. Le script suppose que vous utilisez l'ID utilisateur administrateur par défaut,*jovyan*.

```
#!/bin/bash

CURL="curl --silent -k"
HOST=$(curl -s http://169.254.169.254/latest/meta-data/local-hostname)

admin_token() {
    local user=jovyan
    local pwd=jupyter
    local token=$($CURL https://$HOST:9443/hub/api/authorizations/token \
        -d "{\"username\":\"$user\", \"password\":\"$pwd\"}" | jq ".token")
    if [[ $token != null ]]; then
        token=$(echo $token | sed 's/"//g')
    else
        echo "Unable to get Jupyter API Token."
        exit 1
    fi
    echo $token
}

# Get Jupyter Admin token
token=$(admin_token)
```

```
# Get list of Jupyter users
users=$(curl -XGET -s -k https://$HOST:9443/hub/api/users \
 -H "Authorization: token $token" | jq '.[].name' | sed 's/"//g')

# Create HDFS home dir
for user in ${users[@]};
do
 echo "Create hdfs home dir for $user"
 hadoop fs -mkdir /user/$user
 hadoop fs -chmod 777 /user/$user
done
```

Installation de noyaux et de bibliothèques supplémentaires

Lorsque vous créez un cluster avec JupyterHub Amazon EMR, le noyau Python 3 par défaut pour Jupyter ainsi que les noyaux Spark pour Sparkmagic sont installés sur le PySpark conteneur Docker. Vous pouvez installer des noyaux supplémentaires. Vous pouvez également installer des bibliothèques et des packages supplémentaires, puis les importer pour le shell approprié.

Installation d'un noyau

Les noyaux sont installés dans le conteneur Docker. La manière la plus simple de faire cela consiste à créer un script bash avec des commandes d'installation, de l'enregistrer sur le nœud principal, puis d'utiliser la commande `sudo docker exec jupyterhub` *`script_name`* pour exécuter le script dans le conteneur `jupyterhub`. L'exemple de script suivant installe le noyau, puis installe quelques bibliothèques pour ce noyau sur le nœud principal afin que vous puissiez ultérieurement importer les bibliothèques à l'aide du noyau dans Jupyter.

```
#!/bin/bash

# Install Python 2 kernel
conda create -n py27 python=2.7 anaconda
source /opt/conda/envs/py27/bin/activate
apt-get update
apt-get install -y gcc
/opt/conda/envs/py27/bin/python -m pip install --upgrade ipykernel
/opt/conda/envs/py27/bin/python -m ipykernel install

# Install libraries for Python 2
/opt/conda/envs/py27/bin/pip install paramiko nltk scipy numpy scikit-learn pandas
```

Pour installer le noyau et les bibliothèques dans le conteneur, ouvrez une connexion terminale sur le nœud principal, enregistrez le script sous `/etc/jupyter/install_kernels.sh`, et exécutez la commande suivante sur la ligne de commande du nœud principal :

```
sudo docker exec jupyterhub bash /etc/jupyter/install_kernels.sh
```

Utilisation des bibliothèques et installation de bibliothèques supplémentaires

Un ensemble de bibliothèques de machine learning et de Science des données pour Python 3 est préinstallé JupyterHub sur Amazon EMR. Vous pouvez utiliser `sudo docker exec jupyterhub bash -c "conda list"` et `sudo docker exec jupyterhub bash -c "pip freeze"`.

Si une tâche Spark a besoin de bibliothèques sur les nœuds de travail, nous vous recommandons d'utiliser une action d'amorçage pour exécuter un script afin d'installer les bibliothèques lorsque vous créez le cluster. Les actions d'amorçage s'exécutent sur tous les nœuds du cluster pendant le processus de création du cluster, ce qui simplifie l'installation. Si vous installez des bibliothèques sur des nœuds principaux/de travail une fois que le cluster est en cours d'exécution, l'opération est plus complexe. Dans cette section, nous fournissons un exemple de programme Python qui montre comment installer ces bibliothèques.

Les exemples d'action d'amorçage et de programme Python illustrés dans cette section utilisent un script bash enregistré dans Amazon S3 pour installer les bibliothèques sur tous les nœuds.

Le script référencé dans l'exemple suivant utilise `pip` pour installer paramiko, nltk, scipy, scikit-learn et pandas pour le noyau Python 3 :

```
#!/bin/bash

sudo python3 -m pip install boto3 paramiko nltk scipy scikit-learn pandas
```

Après avoir créé le script, chargez-le dans un emplacement dans Amazon S3 par exemple, `s3://amzn-s3-demo-bucket/install-my-jupyter-libraries.sh`. Pour plus d'informations, consultez la section Chargement d'objets dans le Guide de l'utilisateur Amazon Simple Storage Service afin de pouvoir les utiliser dans votre action de démarrage ou dans votre programme Python.

Pour spécifier une action d'amorçage qui installe des bibliothèques sur tous les nœuds lorsque vous créez un cluster à l'aide du AWS CLI

1. Créez un script semblable à l'exemple précédent et enregistrez-le dans un emplacement dans Amazon S3. Nous utilisons l'exemple `s3://amzn-s3-demo-bucket/install-my-jupyter-libraries.sh`.
2. Créez le cluster avec l'option JupyterHub et utilisez l'`Path`argument de l'`--bootstrap-actions`option pour spécifier l'emplacement du script, comme indiqué dans l'exemple suivant :

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
--applications Name=JupyterHub --log-uri s3://amzn-s3-demo-bucket/
MyJupyterClusterLogs \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
 KeyName=MyKeyPair \
--bootstrap-actions Path=s3://amzn-s3-demo-bucket/install-my-jupyter-
libraries.sh,Name=InstallJupyterLibs
```

Pour spécifier une action d'amorçage qui installe des bibliothèques sur tous les nœuds lorsque vous créez un cluster à l'aide de la console

1. Accédez à la nouvelle console Amazon EMR et sélectionnez Changer pour l'ancienne console depuis le menu latéral. Pour plus d'informations sur ce qu'implique le passage à l'ancienne console, consultez la rubrique Utilisation de l'ancienne console.
2. Choisissez Créer un cluster et Go to advanced options (Aller aux options avancées).
3. Spécifiez les paramètres pour Software and Steps (Logiciel et étapes) et Hardware (Matériel) selon les besoins de votre application.
4. Sur l'écran General Cluster Settings (Paramètres généraux de cluster), développez Bootstrap Actions (Actions d'amorçage).

5. Pour Add bootstrap action (Ajouter une action d'amorçage), sélectionnez Custom action (Action personnalisée), Configure and add (Configurer et ajouter).
6. Pour Name (Nom), saisissez un nom convivial. Pour Emplacement du script, saisissez l'emplacement de votre script dans Amazon S3 (l'exemple que nous utilisons est s3://amzn-s3-demo-bucket/ install-my-jupyter-libraries .sh). Laissez les Optional arguments (Arguments facultatifs) vides et choisissez Add (Ajouter).
7. Spécifiez les autres paramètres de votre cluster, puis choisissez Next (Suivant).
8. Spécifiez les paramètres de sécurité, puis choisissez Create cluster (Créer le cluster).

Example Installation de bibliothèques sur les nœuds principaux d'un cluster en cours d'exécution

Après avoir installé des bibliothèques sur le nœud principal à partir de Jupyter, vous pouvez installer des bibliothèques sur les nœuds principaux en cours d'exécution de diverses manières. L'exemple suivant présente un programme Python écrit pour s'exécuter sur une machine locale. Lorsque vous exécutez le programme Python localement, il utilise le symbole `AWS-RunShellScript` de AWS Systems Manager pour exécuter l'exemple de script, présenté plus haut dans cette section, qui installe des bibliothèques sur les nœuds principaux du cluster.

```
import argparse
import time
import boto3


def install_libraries_on_core_nodes(cluster_id, script_path, emr_client, ssm_client):
    """
    Copies and runs a shell script on the core nodes in the cluster.

    :param cluster_id: The ID of the cluster.
    :param script_path: The path to the script, typically an Amazon S3 object URL.
    :param emr_client: The Boto3 Amazon EMR client.
    :param ssm_client: The Boto3 AWS Systems Manager client.
    """
    core_nodes = emr_client.list_instances(
        ClusterId=cluster_id, InstanceGroupTypes=["CORE"]
    )["Instances"]
    core_instance_ids = [node["Ec2InstanceId"] for node in core_nodes]
    print(f"Found core instances: {core_instance_ids}.")

    commands = [
        # Copy the shell script from Amazon S3 to each node instance.
```

```
        f"aws s3 cp {script_path} /home/hadoop",
        # Run the shell script to install libraries on each node instance.
        "bash /home/hadoop/install_libraries.sh",
    ]
    for command in commands:
        print(f"Sending '{command}' to core instances...")
        command_id = ssm_client.send_command(
            InstanceIds=core_instance_ids,
            DocumentName="AWS-RunShellScript",
            Parameters={"commands": [command]},
            TimeoutSeconds=3600,
        )["Command"]["CommandId"]
        while True:
            # Verify the previous step succeeded before running the next step.
            cmd_result = ssm_client.list_commands(CommandId=command_id)["Commands"][0]
            if cmd_result["StatusDetails"] == "Success":
                print(f"Command succeeded.")
                break
            elif cmd_result["StatusDetails"] in ["Pending", "InProgress"]:
                print(f"Command status is {cmd_result['StatusDetails']}, waiting...")
                time.sleep(10)
            else:
                print(f"Command status is {cmd_result['StatusDetails']}, quitting.")
                raise RuntimeError(
                    f"Command {command} failed to run. "
                    f"Details: {cmd_result['StatusDetails']}"
                )


def main():
    parser = argparse.ArgumentParser()
    parser.add_argument("cluster_id", help="The ID of the cluster.")
    parser.add_argument("script_path", help="The path to the script in Amazon S3.")
    args = parser.parse_args()

    emr_client = boto3.client("emr")
    ssm_client = boto3.client("ssm")

    install_libraries_on_core_nodes(
        args.cluster_id, args.script_path, emr_client, ssm_client
    )


if __name__ == "__main__":
```

```
    main()
```

JupyterHub historique des publications

Le tableau suivant répertorie la version JupyterHub incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

JupyterHub informations sur la version

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-7.9.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.8.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server,

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
		spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7,7.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.6.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-7.5.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.4.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-7.3.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.2.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.36.2	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.1.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-7.0.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.15.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.14.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.13.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.12.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.11.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.11.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.10.1	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.10.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.9.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.9.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.8.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.8.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.7.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.36.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.36.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.6.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.35.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.5.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.4.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.3.1	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.3.0	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.2.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.2.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.1.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.1.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-6.0.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.0.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.34.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.33.1	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.33.0	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.32.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.32.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.31.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.31.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.30.2	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.30.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.30.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
mr-5.29.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.28.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.28.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.27.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.27.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.26.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.25.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.24.1	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.24.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.23.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.23.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.22.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.21.2	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.21.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.21.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.20.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.20.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.19.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.19.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.18.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.18.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.17.2	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.17.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.17.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.16.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.16.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.15.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.15.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.14.2	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.14.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Étiquette de version Amazon EMR	JupyterHub Version	Composants installés avec JupyterHub
emr-5.14.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Apache Livy

Livy permet l'interaction sur une interface REST avec un cluster EMR exécutant Spark. Vous pouvez utiliser l'interface REST ou une bibliothèque client RPC pour soumettre des tâches Spark ou des extraits de code Spark, extraire les résultats de façon synchrone ou asynchrone, et gérer SparkContext. Pour plus d'informations, consultez le site web Apache Livy. Livy est inclus dans Amazon EMR version 5.9.0 ou ultérieure.

Pour accéder à l'interface web Livy, configurez un tunnel SSH sur le nœud principal et une connexion proxy. Pour plus d'informations, consultez Affichage des interfaces web hébergées sur des clusters Amazon EMR.

Le tableau suivant répertorie la version de Livy incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants qu'Amazon EMR installe avec Livy.

Pour connaître la version des composants installés avec Livy dans cette version, consultez les versions des composants de la version 7.9.0.

Informations sur la version de Livy pour emr-7.9.0

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-7.9.0	Livy 0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Le tableau suivant répertorie la version de Livy incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Livy.

Pour connaître la version des composants installés avec Livy dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version de Livy pour emr-6.15.0

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.15.0	Livy 0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Le tableau suivant répertorie la version de Livy incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Livy.

Pour connaître la version des composants installés avec Livy dans cette version, consultez les versions des composants de la version 5.36.2.

Informations sur la version de Livy pour emr-5.36.2

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.36.2	Livy 0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
		yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Rubriques

- Activation du HTTPS avec Apache Livy
- Historique des versions de Livy

Activation du HTTPS avec Apache Livy

Cette rubrique est pertinente si vous utilisez Amazon EMR 7.3.0 ou une version antérieure. À partir de la version 7.4.0, le HTTPS est activé par défaut avec Apache Livy.

1. Provisionnez un cluster Amazon EMR avec le chiffrement du transit activé. Pour en savoir plus sur le chiffrement, consultez Chiffrer les données au repos et en transit.
2. Créez un fichier nommé `livy_ssl.sh` avec le contenu suivant.

```
#!/bin/bash

KEYSTORE_FILE=`awk '/ssl.server.keystore.location/{getline; print}' /etc/hadoop/
conf/ssl-server.xml | sed -e 's/<[^>]*>//g' | tr -d ' \t\n\r\f'`
KEYSTORE_PASS=`awk '/ssl.server.keystore.password/{getline; print}' /etc/hadoop/
conf/ssl-server.xml | sed -e 's/<[^>]*>//g' | tr -d ' \t\n\r\f'`
KEY_PASS=`awk '/ssl.server.keystore.keypassword/{getline; print}' /etc/hadoop/conf/
ssl-server.xml | sed -e 's/<[^>]*>//g' | tr -d ' \t\n\r\f'`

echo "livy.keystore $KEYSTORE_FILE
livy.keystore.password $KEYSTORE_PASS
livy.key-password $KEY_PASS" | sudo tee -a /etc/livy/conf/livy.conf >/dev/null
```

```
sudo systemctl restart livy-server.service
```

3. Exécutez le script suivant en tant qu'étape Amazon EMR. Ce script modifie `/etc/livy/conf/livy.conf` pour activer le protocole SSL.

```
--steps '[{"Args":["s3://amzn-s3-demo-bucket/
livy_ssl.sh"],"Type":"CUSTOM_JAR","ActionOnFailure":"CONTINUE","Jar":"s3://
us-east-1.elasticmapreduce/libs/script-runner/script-
runner.jar","Properties":"","Name":"Custom JAR"}]'
```

4. Redémarrez le service Apache Livy afin que la modification prenne effet. Pour redémarrer Apache Livy, consultez la section Arrêt et redémarrage des processus.
5. Vérifiez que les clients peuvent désormais communiquer via HTTPS. Pour soumettre une tâche, par exemple, exécutez le code suivant.

```
curl -k -X POST --data '{"file": "local:///usr/lib/spark/examples/jars/spark-
examples.jar",
"className": "org.apache.spark.examples.SparkPi"}' \
-H "Content-Type: application/json" \
https://EMR_Master_Node_Host:8998/batches
```

Si vous avez activé HTTPS avec succès, Livy envoie une réponse indiquant que la commande a été acceptée et que le traitement par lots a été soumis.

```
{"id":1,"name":null,"owner":null,"proxyUser":null,"state":"starting","appId":null,"appInfo"
{"driverLogUrl":null,"sparkUiUrl":null},"log":["stdout: ","\nstderr: ","\nYARN
 Diagnostics: "]}
```

Historique des versions de Livy

Le tableau suivant répertorie la version de Livy incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version de Livy

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-7.9.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.8.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7,7.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager,

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
		hadoop-yarn-timeline-server , hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.6.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.5.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-7.4.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.3.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-7.2.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.36.2	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-7.1.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.0.0	0.7.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.15.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.14.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.13.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.12.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.11.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.11.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.10.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.10.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.9.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.9.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.8.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.8.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.7.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.36.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.36.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.6.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.35.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.5.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.4.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.3.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.3.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.2.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.2.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.1.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.1.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.0.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-6.0.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.34.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.33.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.33.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.32.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.32.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.31.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.31.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.30.2	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.30.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.30.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-notebook-env, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
mr-5.29.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.28.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.28.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.27.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.27.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.26.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.25.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.24.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.24.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.23.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.23.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.22.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.21.2	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.21.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.21.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.20.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.20.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.19.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.19.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.18.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.18.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.17.2	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.17.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.17.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.16.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.16.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.15.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.15.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.14.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.14.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.14.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.13.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.13.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.12.3	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.12.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.12.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.12.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.11.4	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.11.3	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.11.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.11.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.11.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.10.1	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.10.0	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.9.1	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Étiquette de version Amazon EMR	Version Livy	Composants installés avec Livy
emr-5.9.0	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Apache MXNet

> **Note**
>
> MXNet est arrivé en fin de vie et le projet est archivé. Amazon EMR 7.2.0 est la dernière version qui inclut. MXNet

Apache MXNet est une bibliothèque d'accélération conçue pour créer des réseaux de neurones et d'autres applications d'apprentissage profond. MXNetautomatise les flux de travail courants et optimise les calculs numériques. MXNet vous aide à concevoir des architectures de réseaux neuronaux sans avoir à vous concentrer sur la mise en œuvre de calculs de bas niveau, tels que les opérations d'algèbre linéaire. MXNet est inclus dans les versions 5.10.0 et ultérieures d'Amazon EMR.

Pour plus d'informations, consultez le site MXNet Web d'Apache.

Le tableau suivant répertorie la version MXNet incluse dans la dernière version de la série Amazon EMR 7.x, ainsi que les composants avec lesquels Amazon EMR est installé. MXNet

Pour connaître la version des composants installés avec MXNet, consultez la version 7.2.0 des versions des composants.

MXNet informations de version pour emr-7.2.0

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-7.2.0	MXNet 1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Le tableau suivant répertorie la version MXNet incluse dans la dernière version de la série Amazon EMR 6.x, ainsi que les composants avec lesquels Amazon EMR est installé. MXNet

Pour connaître la version des composants installés MXNet dans cette version, consultez la version 6.15.0 Versions des composants.

MXNet informations de version pour emr-6.15.0

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-6.15.0	MXNet 1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Le tableau suivant répertorie la version MXNet incluse dans la dernière version de la série Amazon EMR 5.x, ainsi que les composants avec lesquels Amazon EMR est installé. MXNet

Pour la version des composants installés MXNet dans cette version, voir la version 5.36.2 Versions des composants.

MXNet informations de version pour emr-5.36.2

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.36.2	MXNet 1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
		cemanager, hadoop-yarn-timeline-server, mxnet, opencv

MXNet historique des versions

Le tableau suivant répertorie la version MXNet incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

MXNet informations sur la version

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-7.2.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.36.2	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
		cemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-7.1.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-7.0.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.15.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-6.14.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.13.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.12.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-6.11.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.11.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.10.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-6.10.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.9.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.9.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-6.8.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.8.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.7.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.36.1	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.36.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.6.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.35.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.5.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.4.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-6.3.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.3.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.2.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-6.2.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.1.1	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.1.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-6.0.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.0.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.34.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.33.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.33.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.32.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.32.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.31.1	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.31.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.30.2	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.30.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.30.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
mr-5.29.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.28.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.28.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.27.1	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.27.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.26.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.25.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.24.1	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.24.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.23.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.23.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.22.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.21.2	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.21.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.21.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.20.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.20.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.19.1	1.3.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.19.0	1.3.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.18.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.18.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.17.2	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.17.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.17.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.16.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.16.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.15.1	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.15.0	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.14.2	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.14.1	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.14.0	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.13.1	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.13.0	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.12.3	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.12.2	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.12.1	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.12.0	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.4	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.3	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.11.2	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.1	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.0	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Étiquette de version Amazon EMR	MXNet Version	Composants installés avec MXNet
emr-5.10.1	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.10.0	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Apache Oozie

> **Note**
>
> Obsolète d'Oozie — OSS Oozie étant passé en mode attique, il sera supprimé de l'une des prochaines versions d'EMR. Nous recommandons aux clients d'utiliser MWAA comme alternative à Oozie. Vous pouvez vous référer à l'utilitaire pour GitHub migrer oozie-to-airflow-emrles flux de travail Oozie existants vers MWAA.

Utilisez le planificateur de flux de travail Apache Oozie pour gérer et coordonner des travaux Hadoop. Pour de plus amples informations, veuillez consulter https://attic.apache.org/projects/oozie.html.

L'interface web native Oozie n'est pas prise en charge sur Amazon EMR. Pour utiliser une interface frontale pour Oozie, essayez l'application Hue Oozie. Pour de plus amples informations, veuillez consulter Hue. Oozie est inclus avec Amazon EMR version 5.0.0 ou ultérieure. Oozie est inclus sous la forme d'une application sandbox dans les versions précédentes. Pour de plus amples informations, veuillez consulter Versions Amazon EMR 4.x.

Si vous utilisez une AMI Amazon Linux personnalisée basée sur une AMI Amazon Linux dont la date de création est le 11/08/2018, le serveur Oozie ne démarre pas. Si vous utilisez Oozie, créez une AMI personnalisée basée sur un ID d'AMI Amazon Linux avec une date de création différente. Vous pouvez utiliser la AWS CLI commande suivante pour renvoyer une liste d'images IDs pour tous les HVM Amazon Linux AMIs avec une version 2018.03, ainsi que la date de sortie, afin de pouvoir choisir une AMI Amazon Linux appropriée comme base. Remplacez MyRegion par l'identifiant de votre région, tel que us-west-2.

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`]|[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

Le tableau suivant répertorie la version d'Oozie incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants qu'Amazon EMR installe avec Oozie.

Pour connaître la version des composants installés avec Oozie dans cette version, consultez Versions des composants de la version 7.9.0.

Informations sur la version d'Oozie pour emr-7.9.0

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-7.9.0	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Le tableau suivant répertorie la version d'Oozie incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Oozie.

Pour connaître la version des composants installés avec Oozie dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version d’Oozie pour emr-6.15.0

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.15.0	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodema

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
		nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Le tableau suivant répertorie la version d'Oozie incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Oozie.

Pour connaître la version des composants installés avec Oozie dans cette version, consultez Versions des composants de la version 5.36.2.

Informations sur la version d'Oozie pour emr-5.36.2

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.36.2	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Rubriques

- Utilisation d'Oozie avec une base de données distante dans Amazon RDS
- Configurer la version Java pour Oozie

- Historique des versions d'Oozie

Utilisation d'Oozie avec une base de données distante dans Amazon RDS

Par défaut, les informations utilisateur et les historiques de requêtes Oozie sont stockés dans une base de données MySQL locale sur le nœud principal (maître). Vous pouvez également créer un ou plusieurs clusters compatibles avec Oozie à l'aide d'une configuration stockée dans Amazon S3 et d'une base de données MySQL dans Amazon Relational Database Service (Amazon RDS). Cela vous permet de conserver les informations utilisateur et l'historique des requêtes créés par Oozie sans maintenir votre cluster Amazon EMR en cours d'exécution. Nous vous recommandons d'utiliser le chiffrement côté serveur Amazon S3 pour stocker le fichier de configuration.

Commencez par créer la base de données distante pour Oozie.

Pour créer la base de données MySQL externe

1. Ouvrez la console Amazon RDS à l'adresse https://console.aws.amazon.com/rds/.
2. Sélectionnez Launch a DB Instance (Lancer une instance de base de donnéesà.
3. Choisissez MySQL, puis Select (Sélectionner).
4. Laissez la sélection par défaut de Multi-AZ Deployment and Provisioned IOPS Storage (Déploiement multi-AZ et Stockage sur IOPS provisionnés) et choisissez Next (Suivant).
5. Conservez les valeurs par défaut des spécifications de l'instance, spécifiez les paramètres, puis choisissez Next (Suivant).
6. Dans la page de configuration de paramètres avancés, choisissez des noms de groupe de sécurité et de base de données adéquats. Le groupe de sécurité que vous utilisez doit permettre au moins l'accès TCP entrant pour le port 3306 à partir du nœud principal de votre cluster. Si vous n'avez pas créé votre cluster à ce stade, vous pouvez autoriser tous les hôtes à se connecter au port 3306 et ajuster le groupe de sécurité, une fois que vous avez lancé le cluster. Choisissez Launch DB Instance (Lancer l'instance de base de données).
7. Dans le tableau de bord RDS, sélectionnez Instances et choisissez l'instance que vous venez de créer. Lorsque votre base de données est disponible, notez la dbname (nom de la base de données), le nom d'utilisateur, le mot de passe et le nom d'hôte d'instance RDS. Vous allez utiliser ces informations pour créer et configurer votre cluster.

Pour spécifier une base de données MySQL externe pour Oozie lors du lancement d'un cluster à l'aide du AWS CLI

Pour spécifier une base de données MySQL externe pour Oozie lors du lancement d'un cluster à l'aide de l' AWS CLI, utilisez les informations que vous avez notées lors de la création de votre instance RDS pour configurer `oozie-site` avec un objet de configuration.

> **Note**
>
> Vous pouvez créer plusieurs clusters qui utilisent la même base de données externe, mais chaque cluster partage l'historique des requêtes et les informations utilisateur.

- À l'aide de AWS CLI, créez un cluster sur lequel Oozie est installé, en utilisant la base de données externe que vous avez créée et en référençant un fichier de configuration avec une classification de configuration pour Oozie qui spécifie les propriétés de la base de données. L'exemple suivant crée un cluster avec Oozie installé en référençant un fichier de configuration dans Amazon S3, `myConfig.json`, qui spécifie la configuration de la base de données.

  > **Note**
  >
  > Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

  ```
  aws emr create-cluster --release-label emr-7.9.0 --applications Name=Oozie
   Name=Spark Name=Hive \
  --instance-type m5.xlarge --instance-count 3 \
  --configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/
  myConfig.json --use-default-roles
  ```

  Vous trouverez ci-dessous des exemples de contenus du fichier `myConfig.json`. Remplacez *JDBC URLusername*, et *password* par l'URL JDBC, le nom d'utilisateur et le mot de passe de votre instance RDS.

> **Important**
>
> L'URL JDBC doit inclure le nom de la base de données comme suffixe. Par exemple, jdbc:mysql ://.xxxxxx.us-east-1.rds.Amazaws.com:3306/ dbname. oozie-external-db

```
[{
  "Classification": "oozie-site",
    "Properties": {
        "oozie.service.JPAService.jdbc.driver": "org.mariadb.jdbc.Driver",
        "oozie.service.JPAService.jdbc.url": "JDBC URL",

        "oozie.service.JPAService.jdbc.username": "username",
        "oozie.service.JPAService.jdbc.password": "password"
    },
    "Configurations": []
}]
```

Configurer la version Java pour Oozie

Oozie exécute plusieurs processus de machine virtuelle Java (JVM). Cette page explique comment configurer la version Java pour chaque processus.

- Serveur Oozie : définissez `JAVA_HOME` dans la classification `oozie-env` pour mettre à jour la version Java pour `EmbeddedOozieServer`.
- Oozie Launcher AM : Oozie Launcher AM est une tâche MR à mappeur unique qui invoque les bibliothèques clientes d'applications appropriées telles que Hadoop et Hive. Sauf configuration contraire, les versions d'exécution d'Oozie Launcher AM sont les mêmes que celles de Java pour Hadoop dans le cluster EMR. Pour configurer l'environnement d'exécution Java pour Oozie Launcher AM, définissez la propriété suivante dans `workflow.xml` de la tâche :

```
<property>
     <name>mapred.child.env</name>
     <value>JAVA_HOME=/path/to/JAVA_HOME</value>
 </property>
```

Cette propriété garantit que le lanceur Oozie AM pour le job Oozie s'exécute sur la version Java que vous spécifiez, plutôt que sur la version Java définie dans Hadoop.

- Exécutable du client d'application : étant donné qu'Oozie Launcher AM appelle le client de l'application par défaut, le runtime Java de l'exécutable client est le même que celui d'Oozie Launcher AM.
- Applications lancées par une tâche Oozie : sauf indication contraire, les versions d'exécution de l'application en question lancées par une tâche Oozie sont les mêmes JVMs que les environnements d'exécution Java pour Hadoop dans le cluster EMR. Selon le type d'action Oozie Workflow utilisé pour lancer l'application dans un job Oozie (action Spark ou Hive), vous pouvez mettre à jour le runtime Java par défaut pour l'application JVMs en question dans le `workflow.xml` job Oozie.

Historique des versions d'Oozie

Le tableau suivant répertorie la version d'Oozie incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version de Oozie

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-7.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client,

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
		oozie-server, tez-on-yarn, tez-on-worker
emr-7.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7,7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-7.6.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.5.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-7.4.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.3.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-7.2.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-5.36.2	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-7.1.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.0.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.15.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.14.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.13.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.12.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.11.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.11.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.10.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.10.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.9.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.8.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.36.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.36.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.6.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.35.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.5.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.4.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.3.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.3.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.2.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.2.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.1.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.1.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.0.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-6.0.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.34.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.33.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.33.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.32.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.32.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.31.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.31.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.30.2	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.30.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.30.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
mr-5.29.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.28.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.28.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.27.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.27.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.26.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.25.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.24.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.24.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.23.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.23.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.22.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.21.2	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.21.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.21.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.20.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.20.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.19.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.19.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.18.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.18.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.17.2	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.17.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.17.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.16.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.16.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.15.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.15.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.14.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.14.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.14.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.13.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.13.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.12.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.12.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.12.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.12.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.11.4	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.11.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.11.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.11.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.11.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.10.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.10.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.9.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.9.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.8.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.8.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.8.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.8.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.7.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.7.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.6.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.6.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.4	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.5.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.5.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.4.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.4.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.3.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.3.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.3.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.2.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.2.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.2.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.2.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.1.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.1.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.0.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.0.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.0.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Oozie	Composants installés avec Oozie
emr-5.0.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Notes de mise à jour d'Oozie par version

Rubriques

- Amazon EMR 7.6.0 - Modifications apportées à Oozie
- Amazon EMR 7.5.0 - Modifications apportées à Oozie
- Amazon EMR 7.4.0 - Modifications apportées à Oozie
- Amazon EMR 7.3.0 - Modifications apportées à Oozie
- Amazon EMR 7.1.0 - Modifications apportées à Oozie
- Amazon EMR 7.0.0 - Modifications apportées à Oozie

Amazon EMR 7.6.0 - Modifications apportées à Oozie

Aucune modification pour la version 7.6.0.

Amazon EMR 7.5.0 - Modifications apportées à Oozie

Type	Description
Upgrade	Supprimer le support pour l'orchestration des tâches Sqoop car Sqoop n'est pas disponible dans les nouvelles versions d'EMR

Amazon EMR 7.4.0 - Modifications apportées à Oozie

Modifications apportées par Oozie : AUCUNE MODIFICATION

Amazon EMR 7.3.0 - Modifications apportées à Oozie

Type	Description
Upgrade	Guide de mise à niveau vers la version 4.0.
Modification du comportement par défaut	Définit deux propriétés dans `oozie-site.xml` : • `oozie.service.SparkConfigurationService.spark.configurations=*=/etc/spark/conf/` • `oozie.service.SparkConfigurationService.spark.configurations.blacklist=spark.yarn.dist.files` Ces propriétés permettent à Oozie de choisir les configurations Amazon EMR Spark par défaut lorsque vous installez Oozie en même temps que Spark.

Amazon EMR 7.1.0 - Modifications apportées à Oozie

Type	Description
Upgrade	Met à jour Netty vers la version 4.1.100.Final.
Upgrade	Mise à niveau de Jetty vers la version 9.4.53.v20231009.

Amazon EMR 7.0.0 - Modifications apportées à Oozie

Type	Description
Upgrade	Le client et le serveur Oozie utilisent désormais Java 17 par défaut. Pour plus de détails, consultez le guide de mise à jour d'EMR 7.0.0.

Apache Phoenix

Apache Phoenix est utilisé pour l'OLTP et l'analyse opérationnelle, ce qui vous permet d'utiliser des requêtes SQL standard et JDBC APIs pour travailler avec un magasin de sauvegarde Apache HBase . Pour plus d'informations, consultez Phoenix en 15 minutes ou moins. Phoenix est inclus dans Amazon EMR version 4.7.0 ou ultérieure.

Si vous effectuez une mise à niveau depuis une version antérieure d'Amazon EMR vers Amazon EMR version 5.4.0 ou supérieure et utilisez une indexation secondaire, mettez à niveau les index locaux comme décrit dans la documentation Apache Phoenix. Amazon EMR supprime les configurations requises de la classification `hbase-site`, mais les index doivent être repeuplés. La mise à niveau en ligne et hors ligne des index est prise en charge. Les mises à niveau sont par défaut en ligne, ce qui signifie que les index sont remplis de nouveau lors de l'initialisation depuis les clients Phoenix en version 4.8.0 ou supérieure. Pour spécifier des mises à niveau hors ligne, définissez la configuration `phoenix.client.localIndexUpgrade` sur false dans la classification `phoenix-site`, puis lancez SSH sur le nœud maître afin d'exécuter `psql [zookeeper] -1`.

Le tableau suivant répertorie la version de Phoenix incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants qu'Amazon EMR installe avec Phoenix.

Pour connaître la version des composants installés avec Phoenix dans cette version, consultez les versions des composants de la version 7.9.0.

Informations sur la version de Phoenix pour emr-7.9.0

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-7.9.0	Phoenix 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
		timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-queryserver, zookeeper-client, zookeeper-server

Le tableau suivant répertorie la version de Phoenix incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Phoenix.

Pour connaître la version des composants installés avec Phoenix dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version de Phoenix pour emr-6.15.0

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.15.0	Phoenix 5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
		server, zookeeper-client, zookeeper-server

Le tableau suivant répertorie la version de Phoenix incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Phoenix.

Pour connaître la version des composants installés avec Phoenix dans cette version, consultez les versions des composants de la version 5.36.2.

Informations sur la version de Phoenix pour emr-5.36.2

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.36.2	Phoenix 4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Rubriques

- Création d'un cluster avec Phoenix
- Clients Phoenix

- Historique des sorties de Phoenix

Création d'un cluster avec Phoenix

Vous installez Phoenix en choisissant l'application lors de la création d'un cluster sur la console ou en utilisant l' AWS CLI. Les procédures et exemples suivants montrent comment créer un cluster avec Phoenix et HBase. Pour plus d'informations sur la création de clusters à l'aide de la console, y compris les Options avancées, consultez Planifier et configurer des clusters dans le Guide de gestion Amazon EMR.

Pour lancer un cluster avec Phoenix installé à l'aide des Quick Options (Options rapides) pour la création d'un cluster sur la console

1. Ouvrez la console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com
2. Choisissez Create cluster (Créer un cluster) pour utiliser la Quick Create (Création rapide).
3. À côté du champ Configuration des logiciels, choisissez la version la plus récente appropriée à votre application. Phoenix apparaît en tant qu'option uniquement lorsque la Version emr-4.7.0 d'Amazon ou ultérieure est sélectionnée.
4. Pour les applications, choisissez la deuxième option HBase: HBase ***ver*** avec Ganglia***ver***, Hadoop, Hive***ver***, Hue ***verver***, Phoenix et. *ver* ZooKeeper *ver*
5. Sélectionnez d'autres options si nécessaire, puis choisissez Create cluster (Créer le cluster).

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

L'exemple suivant lance un cluster avec Phoenix installé à l'aide des paramètres de configuration par défaut.

Pour lancer un cluster avec Phoenix et HBase utiliser le AWS CLI

- Créez le cluster à l'aide de la commande suivante :

```
aws emr create-cluster --name "Cluster with Phoenix" --release-label emr-7.9.0 \
--applications Name=Phoenix Name=HBase --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

Personnalisation des configurations Phoenix

Lors de la création d'un cluster, vous configurez Phoenix en définissant des valeurs `hbase-site.xml` à l'aide de la classification de configuration `hbase-site`.

Pour plus d'informations, consultez Configuration et réglage dans la documentation Phoenix.

L'exemple suivant illustre l'utilisation d'un fichier JSON stocké dans Amazon S3 pour spécifier la valeur `false` pour la propriété `phoenix.schema.dropMetaData`. Plusieurs propriétés peuvent être spécifies pour une classification unique. Pour de plus amples informations, veuillez consulter Configuration des applications. La commande `create-cluster` référence ensuite le fichier JSON comme paramètre `--configurations`.

Le contenu du fichier JSON enregistré au to /amzn-s3-demo-bucket/myfolder/myconfig format .json est le suivant.

```
[
    {
      "Classification": "hbase-site",
      "Properties": {
        "phoenix.schema.dropMetaData": "false"
      }
    }
  ]
```

La commande `create cluster` qui référence le fichier JSON est illustrée dans l'exemple suivant.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Phoenix \
Name=HBase --instance-type m5.xlarge --instance-count 2 \
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myconfig.json
```

> **Note**
>
> La demande de reconfiguration pour toutes les classifications de configuration Phoenix est uniquement prise en charge dans Amazon EMR version 5.23.0 et ultérieures, et n'est

pas prise en charge dans Amazon EMR version 5.21.0 ou 5.22.0. Pour plus d'informations, consultez Fourniture d'une configuration pour un groupe d'instances dans un cluster en cours d'exécution

Clients Phoenix

Vous vous connectez à Phoenix à l'aide d'un client JDBC construit avec des dépendances complètes ou à l'aide du « client léger » qui utilise le serveur de requêtes Phoenix et ne peut être exécutées que sur un nœud maître d'un cluster (par exemple en utilisant un réacheminement de port SSH, une ligne de commande, une étape, un client SQL, etc.). Lorsque vous utilisez le « gros » client JDBC, il doit toujours avoir accès à tous les nœuds du cluster car il se connecte directement aux HBase services. Le client Phoenix « léger » a uniquement besoin d'un accès au serveur de requête Phoenix dans un port par défaut 8765. Il existe plusieurs scripts au sein de Phoenix qui utilisent ces clients.

Utilisation d'une étape Amazon EMR pour effectuer des requêtes à l'aide de Phoenix

La procédure suivante permet de restaurer un instantané à partir de ces données HBase et de les utiliser pour exécuter une requête Phoenix. Vous pouvez étendre cet exemple ou créer un script qui s'appuie sur les clients Phoenix pour répondre à vos besoins.

1. Créez un cluster avec Phoenix installé, à l'aide de la commande suivante :

```
aws emr create-cluster --name "Cluster with Phoenix" --log-uri s3://amzn-s3-demo-bucket/myLogFolder --release-label emr-7.9.0 \
--applications Name=Phoenix Name=HBase --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

2. Créez, puis chargez les fichiers suivants sur Amazon S3 :

   copySnapshot.sh

```
sudo su hbase -s /bin/sh -c 'hbase snapshot export \
 -D hbase.rootdir=s3://us-east-1.elasticmapreduce.samples/hbase-demo-customer-data/snapshot/ \
-snapshot customer_snapshot1 \
-copy-to hdfs://masterDNSName:8020/user/hbase \
-mappers 2 -chuser hbase -chmod 700'
```

runQuery.sh

```
aws s3 cp s3://amzn-s3-demo-bucket/phoenixQuery.sql /home/hadoop/
/usr/lib/phoenix/bin/sqlline-thin.py http://localhost:8765 /home/hadoop/
phoenixQuery.sql
```

phoenixQuery.sql

> **Note**
>
> Vous ne devez inclure `COLUMN_ENCODED_BYTES=0` dans l'exemple suivant que si vous utilisez les versions 5.26.0 et supérieures d'Amazon EMR.

```
CREATE VIEW "customer" (
pk VARCHAR PRIMARY KEY,
"address"."state" VARCHAR,
"address"."street" VARCHAR,
"address"."city" VARCHAR,
"address"."zip" VARCHAR,
"cc"."number" VARCHAR,
"cc"."expire" VARCHAR,
"cc"."type" VARCHAR,
"contact"."phone" VARCHAR)
COLUMN_ENCODED_BYTES=0;

CREATE INDEX my_index ON "customer" ("customer"."state") INCLUDE("PK",
 "customer"."city", "customer"."expire", "customer"."type");

SELECT "customer"."type" AS credit_card_type, count(*) AS num_customers FROM
 "customer" WHERE "customer"."state" = 'CA' GROUP BY "customer"."type";
```

Utilisez le AWS CLI pour envoyer les fichiers dans le compartiment S3 :

```
aws s3 cp copySnapshot.sh s3://amzn-s3-demo-bucket/
aws s3 cp runQuery.sh s3://amzn-s3-demo-bucket/
aws s3 cp phoenixQuery.sql s3://amzn-s3-demo-bucket/
```

3. Créez une table à l'aide de l'étape suivante soumise au cluster que vous avez créé à l'étape 1 :

createTable.json

```
[
  {
    "Name": "Create HBase Table",
    "Args": ["bash", "-c", "echo $'create \"customer\",\"address\",\"cc\",\"contact
\"' | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file://./createTable.json
```

4. Utilisez `script-runner.jar` pour exécuter le script `copySnapshot.sh` que vous avez téléchargé précédemment sur votre compartiment S3 :

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="HBase Copy Snapshot",ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-
runner.jar,Args=["s3://amzn-s3-demo-bucket/copySnapshot.sh"]
```

   Cela exécute une MapReduce tâche pour copier les données de vos instantanés dans le cluster HDFS.

5. Restaurez l'instantané que vous avez copiée dans le cluster à l'aide de l'étape suivante :

   restoreSnapshot.json

```
[
  {
    "Name": "restore",
    "Args": ["bash", "-c", "echo $'disable \"customer\"; restore_snapshot
 \"customer_snapshot1\"; enable \"customer\"' | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file://./restoreSnapshot.json
```

6. Utilisez `script-runner.jar` pour exécuter le script `runQuery.sh` que vous avez téléchargé précédemment sur votre compartiment S3 :

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Phoenix Run Query",ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-
runner.jar,Args=["s3://amzn-s3-demo-bucket/runQuery.sh"]
```

   La requête s'exécute et renvoie les résultats au `stdout` de l'étape. La réalisation de cette étape peut prendre quelques minutes.

7. Inspectez les résultats du `stdout` de l'étape dans l'URI du journal que vous avez utilisé lorsque vous avez créé le cluster à l'étape 1. Les résultats doivent avoir l'aspect suivant :

```
+-----------------------------------------+-----------------------------------+
|             CREDIT_CARD_TYPE            |              NUM_CUSTOMERS        |
+-----------------------------------------+-----------------------------------+
| american_express                        | 5728                              |
| dankort                                 | 5782                              |
| diners_club                             | 5795                              |
| discover                                | 5715                              |
| forbrugsforeningen                      | 5691                              |
| jcb                                     | 5762                              |
| laser                                   | 5769                              |
| maestro                                 | 5816                              |
| mastercard                              | 5697                              |
| solo                                    | 5586                              |
| switch                                  | 5781                              |
| visa                                    | 5659                              |
+-----------------------------------------+-----------------------------------+
```

Historique des sorties de Phoenix

Le tableau suivant répertorie la version de Phoenix incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des

composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations de version de Phoenix

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-7.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-queryserver, zookeeper-client, zookeeper-server
emr-7.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-reg

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
		ion-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server
emr-7,7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-7.6.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-7.5.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-7.4.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-7.3.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-7.2.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.36.2	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-7.1.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-7.0.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.15.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.14.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.13.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.12.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.11.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.11.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.10.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.10.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.9.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.9.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.8.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.8.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.7.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.36.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.36.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.6.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.35.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.5.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.4.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.3.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.3.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.2.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.2.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.1.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.1.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.0.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-6.0.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.34.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.33.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.33.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.32.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.32.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.31.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.31.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.30.2	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.30.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.30.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
mr-5.29.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.28.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.28.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.27.1	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.27.0	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.26.0	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.25.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.24.1	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.24.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.23.1	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.23.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.22.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.21.2	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.21.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.21.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.20.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.20.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.19.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.19.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.18.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.18.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.17.2	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.17.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.17.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.16.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.16.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.15.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.15.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.14.2	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.14.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.14.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.13.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.13.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.12.3	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.12.2	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.12.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.12.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.11.4	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.11.3	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.11.2	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.11.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.11.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.10.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.10.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.9.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.9.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.8.3	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.8.2	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.8.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.8.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.7.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.7.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.6.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.6.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.4	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.5.3	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.2	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.5.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.4.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.4.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.3.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.3.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.3.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.2.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.2.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.2.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.2.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.1.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.1.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.0.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.0.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.0.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-5.0.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.6	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-4.9.5	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-4.9.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-4.9.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.5	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-4.8.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-4.8.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-4.8.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-4.7.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	Version Phoenix	Composants installés avec Phoenix
emr-4.7.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Apache Pig

Apache Pig est une bibliothèque Apache open source qui s'exécute au-dessus de Hadoop. Elle fournit un langage de script que vous pouvez utiliser pour transformer des ensembles de données volumineux, sans avoir à écrire de code complexe en langage informatique de niveau inférieur, comme Java. La bibliothèque prend des commandes de type SQL écrites dans un langage appelé Pig Latin et les convertit en tâches Tez basées sur des graphes acycliques dirigés (DAGs) ou des programmes. MapReduce Pig fonctionne avec des données structurées et non structurées dans différents formats. Pour plus d'informations sur Pig, consultez la page http://pig.apache.org/.

Vous pouvez exécuter les commandes Pig de façon interactive ou en mode de traitement par lots. Pour utiliser Pig de façon interactive, créez une connexion SSH vers le nœud maître et envoyez des commandes à l'aide du shell Grunt. Pour utiliser Pig en mode de traitement par lots, écrivez vos scripts Pig, chargez-les sur Amazon S3 et envoyez-les en tant qu'étapes du cluster. Pour plus d'informations sur la soumission d'un travail à un cluster, consultez Soumission de travail à un cluster dans le Guide de gestion d'Amazon EMR.

Lorsque vous utilisez Pig pour écrire une sortie dans une HCatalog table dans Amazon S3, désactivez l'écriture directe Amazon EMR en définissant les `mapred.output.direct.EmrFileSystem` propriétés `mapred.output.direct.NativeS3FileSystem` et sur. `false` Pour de plus amples informations, veuillez consulter En utilisant HCatalog. Dans un script Pig, vous pouvez utiliser les commandes `SET mapred.output.direct.NativeS3FileSystem false` et `SET mapred.output.direct.EmrFileSystem false`.

Le tableau suivant répertorie la version de Pig incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants qu'Amazon EMR installe avec Pig.

Pour connaître la version des composants installés avec Pig dans cette version, consultez Versions des composants de la version 7.9.0.

Informations sur la version de Pig pour emr-7.9.0

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-7.9.0	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
		cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Le tableau suivant répertorie la version de Pig incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Pig.

Pour connaître la version des composants installés avec Pig dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version de Pig pour emr-6.15.0

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-6.15.0	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Le tableau suivant répertorie la version de Pig incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Pig.

Pour connaître la version des composants installés avec Pig dans cette version, consultez Versions des composants de la version 5.36.2.

Informations sur la version de Pig pour emr-5.36.2

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.36.2	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Rubriques

- Soumission de travail Pig
- Appel de fonctions définies par l'utilisateur à partir de Pig
- Historique de versions Pig

Soumission de travail Pig

Cette section illustre la soumission d'un travail Pig à un cluster Amazon EMR. Les exemples suivants génèrent un rapport contenant le nombre total d'octets transférés, une liste des 50 adresses IP principales, une liste des 50 référents externes principaux et les 50 critères de recherche principaux sur Bing et Google. Le script Pig est situé dans le compartiment Amazon S3 `s3://elasticmapreduce/samples/pig-apache/do-reports2.pig`. Les données d'entrée sont

situées dans le compartiment Amazon S3 `s3://elasticmapreduce/samples/pig-apache/input`. La sortie est enregistrée dans un compartiment Amazon S3.

Soumettre le travail de Pig à l'aide de la console Amazon EMR

Cet exemple explique comment utiliser la console Amazon EMR pour ajouter une étape Pig à un cluster.

Pour soumettre une étape Pig

1. Ouvrez la console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com
2. Choisissez Créer un cluster pour créer un cluster avec Pig installé. Pour savoir comment créer un cluster, consultez Planifier et configurer un cluster Amazon EMR.
3. Ouvrez un terminal et connectez-vous en SSH au nœud principal de votre cluster en suivant les étapes décrites dans la section Connexion au nœud principal à l'aide de SSH. Une fois que vous avez fait cela, exécutez les étapes suivantes.

   ```
   sudo mkdir -p /home/hadoop/lib/pig/
   sudo aws s3 cp s3://elasticmapreduce/libs/pig/0.3/piggybank-0.3-amzn.jar /home/hadoop/lib/pig/piggybank.jar
   ```

4. Dans la console, cliquez sur Liste des clusters et sélectionnez le nom du cluster que vous avez créé.
5. Faites défiler l'affichage jusqu'à la section Étapes et développez-la, puis choisissez Ajouter une étape.
6. Dans la boîte de dialogue Ajouter une étape :

   - Pour Type d'étape, choisissez Programme Pig.
   - Pour Nom, acceptez le nom par défaut (Programme Pig) ou saisissez un nouveau nom.
   - Pour Emplacement S3 du script, saisissez l'emplacement du script Pig. Par exemple : **`s3://elasticmapreduce/samples/pig-apache/do-reports2.pig`**.
   - Pour Emplacement S3 d'entrée, saisissez l'emplacement des données d'entrée. Par exemple : **`s3://elasticmapreduce/samples/pig-apache/input`**.
   - Pour Emplacement S3 de sortie, saisissez le nom du compartiment de sortie Amazon S3 ou naviguez à ce dernier.
   - Pour Arguments, laissez le champ vide.
   - Pour Action sur échec, acceptez l'option par défaut (Continuer).

7. Choisissez Ajouter. L'étape s'affiche dans la console avec le statut En suspens.
8. Le statut de l'étape passe de Pending (En suspens) à Running (En cours d'exécution) puis à Completed (Terminé), au fur et à mesure de son exécution. Pour mettre à jour l'état, choisissez l'icône d'Actualisation au-dessus de la colonne Actions. Lorsque votre étape est terminée, vérifiez votre compartiment Amazon S3 pour confirmer que les fichiers de sortie de votre étape Pig s'y trouvent.

Soumission d'un travail Pig à l'aide d' AWS CLI

Pour soumettre une étape Pig à l'aide du AWS CLI

Lorsque vous lancez un cluster à l'aide du paramètre AWS CLI, utilisez le `--applications` paramètre pour installer Pig. Pour soumettre une étape Pig, utilisez le paramètre `--steps`.

1. Pour lancer un cluster sur lequel Pig est installé, saisissez la commande suivante, en *amzn-s3-demo-bucket/* remplaçant *myKey* et par le nom de votre paire de EC2 clés et de votre compartiment Amazon S3.

```
aws emr create-cluster \
--name "Test cluster" \
--log-uri s3://amzn-s3-demo-bucket/ \
--release-label emr-5.36.2 \
--applications Name=Pig \
--use-default-roles \
--ec2-attributes KeyName=myKey \
--instance-type m5.xlarge \
--instance-count 3
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

Lorsque vous spécifiez le nombre d'instances sans utiliser le paramètre `--instance-groups`, un seul nœud maître est lancé et les instances restantes sont lancées en tant que nœuds principaux. Tous les nœuds utilisent le type d'instance spécifié dans la commande.

> **Note**
>
> Si vous n'avez pas encore créé le rôle de service EMR par défaut et le profil d' EC2 instance, tapez `aws emr create-default-roles` pour les créer avant de taper la `create-cluster` sous-commande.

2. Pour soumettre une étape Pig, entrez la commande suivante, en *amzn-s3-demo-bucket* remplaçant *myClusterId* et par votre ID de cluster et le nom de votre compartiment Amazon S3.

```
aws emr add-steps \
--cluster-id myClusterId \
--steps Type=PIG,Name="Pig Program",ActionOnFailure=CONTINUE,Args=[-f,s3://
elasticmapreduce/samples/pig-apache/do-reports2.pig,-p,INPUT=s3://elasticmapreduce/
samples/pig-apache/input,-p,OUTPUT=s3://amzn-s3-demo-bucket/pig-apache/output]
```

   Cette commande renverra un identifiant d'étape, que vous pourrez utiliser pour vérifier le `State` de votre étape.

3. Vérifiez l'état de votre étape à l'aide de la commande `describe-step`.

```
aws emr describe-step --cluster-id myClusterId --step-id s-1XXXXXXXXXXA
```

   La valeur `State` de l'étape passe de `PENDING` à `RUNNING`, puis à `COMPLETED`, au fur et à mesure que l'étape s'exécute. Lorsque votre étape est terminée, vérifiez votre compartiment Amazon S3 pour confirmer que les fichiers de sortie de votre étape Pig s'y trouvent.

Pour plus d'informations sur l'utilisation des commandes Amazon EMR dans l'interface AWS CLI, consultez Référence des AWS CLI commandes.

Appel de fonctions définies par l'utilisateur à partir de Pig

Pig permet d'appeler des fonctions définies par l'utilisateur (UDFs) depuis les scripts Pig. Vous pouvez le faire pour implémenter un traitement personnalisé à utiliser dans vos scripts Pig. Les langages actuellement pris en charge sont Java, Python/Jython et JavaScript (bien que le JavaScript support soit encore expérimental).

Les sections suivantes expliquent comment enregistrer vos fonctions avec Pig afin de pouvoir les appeler à partir du shell Pig ou de scripts Pig. Pour plus d'informations sur l'utilisation UDFs avec Pig, consultez la Documentation Pig correspondant à votre version de Pig.

Appel de fichiers JAR à partir de Pig

Vous pouvez utiliser des fichiers JAR personnalisés avec Pig à l'aide de la commande `REGISTER` dans votre script Pig. Le fichier JAR est un système de fichiers local ou un système de fichiers à distance, tel qu'Amazon S3. Lorsque le script Pig s'exécute, Amazon EMR charge le fichier JAR automatiquement vers le nœud principal, puis charge le fichier JAR sur le cache distribué Hadoop. De cette manière, le fichier JAR est automatiquement utilisé selon les besoins par toutes les instances du cluster.

Pour utiliser les fichiers JAR avec Pig

1. Chargez votre fichier JAR personnalisé dans Amazon S3.
2. Utilisez la commande `REGISTER` dans votre script Pig pour spécifier le compartiment sur Amazon S3 du fichier JAR personnalisé.

```
REGISTER s3://amzn-s3-demo-bucket/path/mycustomjar.jar;
```

Appel de scripts Python/Jython à partir de Pig

Vous pouvez enregistrer des scripts Python avec Pig, puis appeler des fonctions dans ces scripts à partir du shell Pig ou dans un script Pig. Pour cela, vous devez spécifier l'emplacement du script avec le mot clé `register`.

Etant donné que Pig est écrit en Java, il utilise le moteur de script Jython pour analyser les scripts Python. Pour plus d'informations sur Jython, accédez à http://www.jython.org/.

Pour appeler un script Python/Jython à partir de Pig

1. Écrivez un script Python et chargez-le dans un emplacement sur Amazon S3. Il doit s'agir d'un compartiment appartenant au même compte que celui qui crée le cluster Pig, ou qui dispose des autorisations définies pour que le compte qui a créé le cluster puisse y accéder. Dans cet exemple, le script est chargé sur s3://amzn-s3-demo -bucket/pig/python.
2. Démarrez un cluster Pig. Si vous accédez à Pig à partir du shell Grunt, exécutez un cluster interactif. Si vous exécutez les commandes Pig à partir d'un script, démarrez un cluster Pig

scripté. Cet exemple démarre un cluster interactif. Pour plus d'informations sur la création d'un cluster Pig, consultez Soumission de travail Pig.

3. Dans le cas d'un cluster interactif, utilisez SSH pour vous connecter au nœud maître et exécutez le shell Grunt. Pour plus d'informations, consultez SSH dans le nœud principal.
4. Exécutez le shell Grunt pour Pig en tapant `pig` dans la ligne de commande :

```
pig
```

5. Enregistrez la bibliothèque Jython et votre script Python avec Pig à l'aide du mot clé `register` à l'invite de commande Grunt, comme illustré dans la commande suivante, dans laquelle vous devez spécifier l'emplacement de votre script dans Amazon S3 :

```
grunt> register 'lib/jython.jar';
grunt> register 's3://amzn-s3-demo-bucket/pig/python/myscript.py' using jython as
 myfunctions;
```

6. Chargez les données d'entrée. L'exemple suivant charge l'entrée à partir d'un emplacement Amazon S3 :

```
grunt> input = load 's3://amzn-s3-demo-bucket/input/data.txt' using TextLoader as
 (line:chararray);
```

7. Vous pouvez désormais appeler des fonctions dans votre script à partir de Pig en y faisant référence à l'aide de `myfunctions` :

```
grunt> output=foreach input generate myfunctions.myfunction($1);
```

Historique de versions Pig

Le tableau suivant répertorie la version de Pig incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version de Pig

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-7.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.8.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7,7.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.6.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.5.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-7.4.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.3.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-7.2.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-5.36.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-7.1.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.0.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-6.15.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.14.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-6.13.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.12.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-6.11.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.11.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-6.10.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.10.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-6.9.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-6.8.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.8.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-6.7.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.36.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.36.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.6.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.35.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.5.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-6.4.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.3.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-6.3.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.2.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-6.2.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.1.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-6.1.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.34.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.33.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.33.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.32.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.32.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.31.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.31.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.30.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.30.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.30.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
mr-5.29.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.28.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.28.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.27.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.27.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.26.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.25.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.24.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.24.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.23.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.23.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.22.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.21.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.21.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.21.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.20.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.20.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.19.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.19.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.18.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.18.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.17.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.17.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.17.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.16.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.16.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.15.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.15.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.14.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.14.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.14.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.13.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.13.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.12.3	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.12.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.12.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.12.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.11.4	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.11.3	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.11.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.11.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.11.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.10.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.10.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.9.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.8.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.8.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.8.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.8.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.7.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.7.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.6.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.6.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.4	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.5.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.5.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.4.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.4.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.3.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.3.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.3.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.2.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.2.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.2.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.2.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.1.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.1.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.0.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.0.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.0.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-5.0.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-4.9.6	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.5	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-4.9.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-4.9.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.5	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-4.8.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-4.8.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-4.7.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-4.6.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.6.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.5.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-4.4.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.3.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.2.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Étiquette de version Amazon EMR	Version de Pig	Composants installés avec Pig
emr-4.1.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.0.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Presto

> **Important**
>
> Presto, la version précédente de Trino, est toujours disponible pour une utilisation avec Amazon EMR. Cependant, nous recommandons vivement d'utiliser Trino à l'avenir avec Amazon EMR. Notez également que Trino et Presto ne peuvent pas fonctionner simultanément sur le même cluster. Pour plus d'informations, consultez Trino.

Presto est un moteur de requête SQL rapide conçu pour exécuter des requêtes analytiques interactives sur des jeux de données volumineux provenant de plusieurs sources. Pour plus d'informations, consultez le site web Presto. Presto est inclus dans Amazon EMR versions 5.0.0 et ultérieures. Les versions antérieures incluent Presto en tant qu'application d'environnement de test (sandbox). Pour de plus amples informations, veuillez consulter Versions Amazon EMR 4.x. Les versions 6.1.0 et ultérieures d'Amazon EMR prennent en charge Trino en plus de Presto. Pour de plus amples informations, veuillez consulter Installation de Prestodb.

Le tableau suivant répertorie la version de Presto incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants qu'Amazon EMR installe avec Presto.

Pour connaître la version des composants installés avec Presto dans cette version, consultez Versions des composants de la version 7.9.0.

Informations sur la version Presto pour emr-7.9.0

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-7.9.0	Presto 0.287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
		client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Le tableau suivant répertorie la version de Presto incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Presto.

Pour connaître la version des composants installés avec Presto dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version Presto pour emr-6.15.0

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.15.0	Presto 0.283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Le tableau suivant répertorie la version de Presto incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Presto.

Pour connaître la version des composants installés avec Presto dans cette version, consultez Versions des composants de la version 5.36.2.

Informations sur la version Presto pour emr-5.36.2

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.36.2	Presto 0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Rubriques

- Utilisation de Presto avec le catalogue de données AWS Glue
- Utilisation de S3 Select Pushdown avec Presto pour améliorer les performances
- Ajout de connecteurs de base de données
- Utilisation du protocole SSL/TLS et configuration de LDAPS avec Presto sur Amazon EMR
- Activation du mode strict Presto
- Gestion de la perte d'instances Spot dans Presto
- Utilisation du dimensionnement automatique de Presto avec désaffectation gracieuse
- Considérations relatives à Presto sur Amazon EMR
- Historique des versions de Presto

Utilisation de Presto avec le catalogue de données AWS Glue

En utilisant les versions 5.10.0 et ultérieures d'Amazon EMR, vous pouvez spécifier le catalogue de données AWS Glue comme métastore Hive par défaut pour Presto. Cette configuration est

recommandée si vous avez besoin d'un metastore persistant ou d'un metastore partagé par différents clusters, services, applications ou Comptes AWS.

AWS Glue est un service d'extraction, transformation et chargement (ETL) entièrement géré qui facilite et rend abordable le classement par catégorie, le nettoyage, l'enrichissement et le déplacement des données de manière fiable entre différentes banques de données. Le catalogue de données AWS Glue fournit un référentiel de métadonnées unifié pour une variété de sources et de formats de données, s'intégrant à Amazon EMR ainsi qu'à Amazon RDS, Amazon Redshift, Redshift Spectrum, Athena, et toute application compatible avec le métastore Apache Hive. AWS Les crawlers Glue peuvent automatiquement déduire le schéma des données sources dans Amazon S3 et stocker les métadonnées associées dans le catalogue de données. Pour plus d'informations sur le catalogue de données, consultez Remplissage du catalogue de données AWS Glue dans le Guide du développeur AWS Glue.

Des frais distincts s'appliquent pour AWS Glue. Il y a un tarif mensuel pour le stockage et l'accès aux métadonnées dans le catalogue de données, un tarif horaire facturé à la minute pour les tâches ETL d' AWS Glue et l'exécution du analyseur, et un prix horaire facturé à la minute pour chaque point de terminaison de développement provisionné. Le catalogue de données vous permet de stocker jusqu'à un million d'objets sans aucuns frais. Au-delà d'un million d'objets, des frais de 1 USD vous seront facturés tous les 100 000 objets. Dans le catalogue de données, sont considérés comme des objets une table, une partition ou une base de données. Pour de plus amples informations, consultez Tarification Glue.

> **Important**
>
> Si vous avez créé des tables à l'aide d'Amazon Athena ou d'Amazon Redshift Spectrum avant le 14 août 2017, les bases de données et les tables sont stockées dans un catalogue géré par Athena, qui est distinct du catalogue de données Glue. AWS Pour intégrer Amazon EMR à ces tables, vous devez mettre à niveau vers le catalogue de données AWS Glue. Pour plus d'informations, consultez la section Mise à niveau vers le catalogue de données AWS Glue dans le guide de l'utilisateur Amazon Athena.

Spécifier le catalogue de données AWS Glue comme métastore

Vous pouvez spécifier le catalogue de données AWS Glue comme métastore à l'aide de la AWS Management Console AWS CLI, de l'ou de l'API Amazon EMR. Lorsque vous utilisez la CLI ou l'API, vous utilisez la classification de configuration de Presto pour spécifier le catalogue de données. En

outre, avec Amazon EMR 5.16.0 et les versions ultérieures, vous pouvez utiliser la classification de configuration pour spécifier un catalogue de données dans un différent. Compte AWS Lorsque vous utilisez la console, vous pouvez spécifier le catalogue de données à l'aide des Options avancées ou des Options rapides.

Console

Pour spécifier le catalogue de données AWS Glue comme métastore Hive dans la nouvelle console

1. Connectez-vous à la et ouvrez AWS Management Console la console Amazon EMR à https://console.aws.amazon.com l'adresse /emr.
2. Sous EMR activé EC2 dans le volet de navigation de gauche, choisissez Clusters, puis Create cluster.
3. Sous Solution groupée d'application, choisissez Presto.
4. Sous Paramètres du catalogue de données AWS Glue, cochez la case Utiliser pour les métadonnées des tables Presto.
5. Choisissez toutes les autres options qui s'appliquent à votre cluster.
6. Pour lancer cluster, choisissez Créer un cluster.

CLI

Pour spécifier le catalogue de données AWS Glue comme métastore Hive par défaut à l'aide de la AWS CLI

Pour des exemples de spécification des classifications de configuration suivantes lorsque vous créez un cluster, consultez Configuration des applications.

Amazon EMR 5.16.0 et versions ultérieures

- Définissez la propriété `hive.metastore` sur `glue`, comme indiqué dans l'exemple JSON suivant.

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore": "glue"
    }
```

```
  }
]
```

Pour spécifier un catalogue de données dans un différent Compte AWS, ajoutez la `hive.metastore.glue.catalogid` propriété, tel qu'illustré dans l'exemple JSON suivant. Remplacez *`acct-id`* par le compte AWS du catalogue de données. L'utilisation d'un catalogue de données dans un autre Compte AWS n'est pas disponible avec Amazon EMR version 5.15.0 et antérieures.

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore": "glue",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

Amazon EMR versions 5.10.0 à 5.15.0

Définissez la propriété `hive.metastore.glue.datacatalog.enabled` sur `true`, comme indiqué dans l'exemple JSON suivant :

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore.glue.datacatalog.enabled": "true"
    }
  }
]
```

Amazon EMR 6.1.0 et versions ultérieures utilisant PrestoSQL (Trino)

À partir de la version 6.1.0 d'EMR, PrestoSQL prend également en charge Glue comme métastore Hive par défaut. Utilisez la classification de configuration `prestosql-connector-hive` et définissez la propriété `hive.metastore` sur `glue`, comme indiqué dans l'exemple JSON suivant.

Les versions 6.4.0 et ultérieures d'Amazon EMR utilisent le nouveau nom Trino au lieu de PrestoSQL. Si vous utilisez Trino, remplacez *prestosql-connector-hive* dans la classification de configuration suivante par `trino-connector-hive`.

```
[
  {
    "Classification": "prestosql-connector-hive",
    "Properties": {
      "hive.metastore": "glue"
    }
  }
]
```

Pour basculer les metastores sur un cluster de longue durée, vous pouvez définir manuellement ces valeurs selon les besoins de votre version en vous connectant au nœud maître, en modifiant les valeurs de propriété directement dans le fichier `/etc/presto/conf/catalog/hive.properties` et en redémarrant le serveur Presto (`sudo restart presto-server`). Si vous utilisez cette méthode sur les versions 5.15.0 et antérieures d'Amazon EMR, assurez-vous que `hive.table-statistics-enabled` soit défini sur `false`. Ce paramètre n'est pas obligatoire lorsque vous utilisez les versions 5.16.0 et ultérieures ; cependant, les statistiques de table et de partition ne sont pas prises en charge.

Autorisations IAM

Le profil d' EC2 instance pour un cluster doit avoir des autorisations IAM pour les actions AWS Glue. En outre, si vous activez le chiffrement pour les objets du catalogue de données AWS Glue, le rôle doit également être autorisé à chiffrer, déchiffrer et générer la AWS KMS key utilisée pour le chiffrement.

Autorisations pour les actions AWS Glue

Aucune action n'est requise si vous utilisez le profil d' EC2 instance par défaut pour Amazon EMR. La politique `AmazonElasticMapReduceforEC2Role` gérée attachée au `EMR_EC2_DefaultRole` autorise toutes les actions AWS Glue nécessaires. Toutefois, si vous spécifiez un profil d' EC2 instance et des autorisations personnalisés, vous devez configurer les actions AWS Glue appropriées. Utilisez la stratégie gérée `AmazonElasticMapReduceforEC2Role` comme point

de départ. Pour plus d'informations, consultez la section Rôle de service pour les EC2 instances de cluster (profil d'EC2 instance) dans le guide de gestion Amazon EMR.

Autorisations de chiffrement et de déchiffrement du catalogue de données Glue AWS

Votre profil d'instance a besoin d'autorisations pour chiffrer et déchiffrer les données au moyen de votre clé. Vous n'avez pas besoin de configurer ces autorisations si les deux déclarations suivantes s'appliquent :

- Vous activez le chiffrement des objets du catalogue de données AWS Glue à l'aide des clés gérées pour AWS Glue.
- Vous utilisez un cluster identique au catalogue Compte AWS de données AWS Glue.

Sinon, vous devez ajouter l'instruction suivante à la politique d'autorisations attachée à votre profil d' EC2 instance.

```
[
    {
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "kms:Decrypt",
                    "kms:Encrypt",
                    "kms:GenerateDataKey"
                ],
                "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
            }
        ]
    }
]
```

Pour plus d'informations sur le chiffrement du catalogue de données AWS Glue, consultez Chiffrement de votre catalogue de données dans le Guide du développeur AWS Glue.

Autorisations basées sur les ressources

Si vous utilisez AWS Glue en conjonction avec Hive, Spark ou Presto dans Amazon EMR, AWS Glue prend en charge les politiques basées sur les ressources pour contrôler l'accès aux ressources

du catalogue de données. Ces ressources comprennent les bases de données, les tables, les connexions et les fonctions définies par l'utilisateur. Pour plus d'informations, consultez Politiques de ressources AWS Glue dans le Guide du développeur AWS Glue.

Lorsque vous utilisez des politiques basées sur les ressources pour limiter l'accès à AWS Glue depuis Amazon EMR, le principal que vous spécifiez dans la politique d'autorisations doit être l'ARN de rôle associé au profil d' EC2 instance qui est spécifié lors de la création d'un cluster. Par exemple, pour une politique basée sur les ressources attachée à un catalogue, vous pouvez spécifier l'ARN de rôle pour le rôle de service par défaut des EC2 instances de cluster, *EMR_EC2_DefaultRole* en utilisant le format indiqué dans l'exemple suivant : `Principal`

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

L'ID de compte AWS Glue *acct-id* peut être différent. Cela permet d'accéder aux clusters EMR à partir de comptes différents. Vous pouvez spécifier plusieurs principaux, chacun provenant d'un compte différent.

Considérations relatives à l'utilisation du catalogue de données AWS Glue

Considérez les éléments suivants lorsque vous utilisez le catalogue de données AWS Glue en tant que métastore avec Presto :

- Renommer des tables à partir d' AWS Glue n'est pas pris en charge.
- Lorsque vous créez une table Hive sans spécifier de `LOCATION`, les données de la table sont stockées à l'emplacement spécifié par la propriété `hive.metastore.warehouse.dir`. Par défaut, il s'agit d'un emplacement dans HDFS. Si un autre cluster doit accéder à la table, il échoue, sauf s'il dispose d'autorisations adéquates sur le cluster qui a créé la table. En outre, le stockage dans HDFS étant transitoire, si le cluster est résilié, les données de la table sont perdues et la table doit être recréée. Nous vous recommandons de spécifier un `LOCATION` dans Amazon S3 lorsque vous créez une table Hive à l'aide d' AWS Glue. Vous pouvez également utiliser la classification de configuration `hive-site` pour spécifier un emplacement dans Amazon S3 pour `hive.metastore.warehouse.dir`, qui s'applique à toutes les tables Hive. Si une table est créée dans un emplacement HDFS et que le cluster qui l'a créée est toujours en cours d'exécution, vous pouvez mettre à jour l'emplacement de la table vers Amazon S3 depuis AWS Glue. Pour plus d'informations, consultez Utilisation des tables dans la console AWS Glue dans le Guide du développeur AWS Glue.
- Les valeurs de partition contenant des guillemets et des apostrophes ne sont pas prises en charge, par exemple `PARTITION (owner="Doe's").`.

- Les statistiques sur les colonnes sont prises en charge à partir de la version emr-5.31.0.
- L'utilisation de l'autorisation Hive n'est pas prise en charge. Comme alternative, envisagez d'utiliser les politiques basées sur les ressources AWS Glue. Pour plus d'informations, consultez Utilisation de politiques basées sur des ressources pour l'accès d'Amazon EMR au catalogue de données Glue AWS.

Utilisation de S3 Select Pushdown avec Presto pour améliorer les performances

> **Important**
>
> Amazon S3 Select n'est plus disponible pour les nouveaux clients. Les clients existants d'Amazon S3 Select peuvent continuer à utiliser cette fonctionnalité comme d'habitude. En savoir plus

Avec la version 5.18.0 et ultérieure d'Amazon EMR, vous pouvez utiliser S3 select Pushdown avec Presto sur Amazon EMR. Cette fonctionnalité permet à Presto de « pousser vers le bas » le travail de calcul d'opérations de projection (par exemple, `SELECT`) et les opérations de prédicat (par exemple, `WHERE`) sur Amazon S3. Cela autorise des requêtes à récupérer uniquement les données nécessaires d'Amazon S3 , ce qui peut améliorer les performances et réduire les quantités de données transférées entre Amazon EMR et Amazon S3 dans certaines applications.

S3 Select Pushdown est-il adapté à mon application ?

Nous vous recommandons de tester vos applications avec et sans S3 Select Pushdown pour voir si son utilisation peut être adaptée à votre application.

Utilisez les consignes suivantes pour déterminer si votre application peut utiliser S3 Select :

- Votre requête exclut plus de la moitié de l'ensemble de données d'origine.
- Vos prédicats de filtre de requête utilisent les colonnes qui ont un type de données pris en charge par Presto et S3 Select. Les types de données horodateur, réel et double ne sont pas pris en charge par S3 Select Pushdown. Nous recommandons d'utiliser le type de données décimal pour les données quantitatives. Pour plus d'informations sur les types de données pris en charge pour S3 Select, veuillez consulter Amazon Simple Storage Service, veuillez consulter Types de données dans le Guide de l'utilisateur Amazon Simple Storage Service.

- Votre connexion réseau entre Amazon S3 et le cluster Amazon EMR a une bonne vitesse de transfert et une bande passante disponible. Amazon S3 ne compresse pas les réponses HTTP, de sorte que la taille de la réponse est susceptible d'augmenter pour les fichiers d'entrée compressés.

Considérations et restrictions

- Seuls les objets stockés en format CSV sont pris en charge. Les objets peuvent être décompressés ou facultativement compressés avec gzip ou bzip2.
- La propriété `AllowQuotedRecordDelimiters` n'est pas prise en charge. Si cette propriété est spécifiée, la requête échoue.
- Le chiffrement côté serveur d'Amazon S3 avec des clés de chiffrement fournies par le client (SSE-C) et le chiffrement côté client ne sont pas pris en charge.
- S3 Select Pushdown n'est pas un substitut pour l'utilisation de formats en colonne ou compressés tels qu'ORC ou Parquet.
- Amazon S3 Select ne prend pas en charge les requêtes entre régions pour Presto et Trino.

Activation de S3 Select Pushdown avec PrestoDB ou Trino

Pour autoriser S3 Select Pushdown pour PrestoDB sur Amazon EMR, utilisez la classification de configuration `presto-connector-hive` pour définir `hive.s3select-pushdown.enabled` sur `true`, tel qu'illustré dans l'exemple suivant. Pour de plus amples informations, veuillez consulter Configuration des applications. La valeur hive.s3select-pushdown.max-connections doit également être définie. Le paramètre *`500`* par défaut devrait être approprié pour la plupart des applications. Pour plus d'informations, consultez Comprendre et régler hive.s3select-pushdown.max-connections ci-dessous.

Pour PrestoSQL sur EMR versions 6.1.0 à 6.3.0, remplacez *`presto-connector-hive`* dans l'exemple ci-dessous par `prestosql-connector-hive`.

Les versions 6.4.0 et ultérieures d'Amazon EMR utilisent le nouveau nom Trino au lieu de PrestoSQL. Si vous utilisez Trino, remplacez *`presto-connector-hive`* dans l'exemple ci-dessous par `trino-connector-hive`

```
[
    {
        "classification": "presto-connector-hive",
```

```
        "properties": {
            "hive.s3select-pushdown.enabled": "true",
            "hive.s3select-pushdown.max-connections": "500"
        }
    }
]
```

Comprendre et régler hive.s3select-pushdown.max-connections

Presto utilise EMRFS comme système de fichier par défaut. Le paramètre `fs.s3.maxConnections` dans la configuration de classification `emrfs-site` spécifie le nombre de connexions de client permissibles à Amazon S3 via EMRFS pour Presto. 500 par défaut. S3 Select Pushdown contourne EMRFS lorsqu'il accède à Amazon S3 pour les opérations de prédicats. Dans ce cas, la valeur de `hive.s3select-pushdown.max-connections` détermine le nombre maximal de connexions clients autorisées pour ces opérations de nœuds de travail. Cependant, toute demande à Amazon S3 initiée par Presto qui n'est pas poussée vers le bas (par exemple, les opérations GET) continue d'être régie par la valeur de `fs.s3.maxConnections`.

Si votre application fait l'expérience de l'erreur « Temporisation attente de connexion du groupe », augmentez la valeur de `hive.s3select-pushdown.max-connections` et de `fs.s3.maxConnections`.

Ajout de connecteurs de base de données

Vous pouvez utilisez les classifications de configuration pour configurer les propriétés de connecteur JDBC lorsque vous créez un cluster. Les classifications de configuration commencent par `presto-connector`, par exemple `presto-connector-postgresql`. Les classifications de configuration disponibles dépendent de la version d'Amazon EMR. Pour les classifications de configuration disponibles avec la version la plus récente, consultez the section called “5.36.2 classifications de configuration” pour Amazon EMR 5.36.2. Si vous utilisez une version différente d'Amazon EMR, consultez Versions Amazon EMR 5.x pour les classifications de configurations. Pour plus d'informations sur les propriétés qui peuvent être configurées avec chaque connecteur, consultez https://prestodb.io/docs/current/connector.html.

Example – Configuration d'un cluster avec le connecteur JDBC PostgreSQL

Pour lancer un cluster avec le connecteur PostgreSQL installé et configuré, créez d'abord un fichier JSON qui spécifie la classification de configuration (par exemple, `myConfig.json`) avec le contenu suivant, puis enregistrez-le localement.

Remplacez les propriétés de connexion de manière adaptée à votre configuration, tel qu'illustré dans la rubrique Connecteur PostgreSQL de la Documentation Presto.

```
[
  {
    "Classification": "presto-connector-postgresql",
    "Properties": {
      "connection-url": "jdbc:postgresql://example.net:5432/database",
      "connection-user": "MYUSER",
      "connection-password": "MYPASS"
    },
    "Configurations": []
  }
]
```

Lorsque vous créez le cluster, faites référence au chemin du fichier JSON qui utilise l'option `--configurations` affichée dans l'exemple suivant, où `myConfig.json` est dans le même répertoire où vous exécutez la commande :

```
aws emr create-cluster --name PrestoConnector --release-label emr-5.36.2 --instance-
type m5.xlarge \
--instance-count 2 --applications Name=Hadoop Name=Hive Name=Pig Name=Presto \
--use-default-roles --ec2-attributes KeyName=myKey \
--log-uri s3://amzn-s3-demo-bucket/logs --enable-debugging \
--configurations file://myConfig.json
```

Utilisation du protocole SSL/TLS et configuration de LDAPS avec Presto sur Amazon EMR

Avec les versions 5.6.0 et ultérieures d'Amazon EMR, vous pouvez activer SSL/TLS pour contribuer à sécuriser les communications internes entre les nœuds Presto. Pour ce faire, vous devez définir une configuration de sécurité pour le chiffrement en transit. Pour plus d'informations, consultez les sections Options de chiffrement et Utiliser les configurations de sécurité pour configurer la sécurité du cluster dans le Guide de gestion Amazon EMR.

Lorsque vous utilisez une configuration de sécurité avec le chiffrement en transit, Amazon EMR exécute les actions suivantes pour Presto :

- Il distribue les artefacts de chiffrement, ou certificats, que vous spécifiez pour le chiffrement en transit sur l'ensemble du cluster Presto. Pour plus d'informations, consultez Mise à disposition des certificats de chiffrement des données en transit.
- Il définit les propriétés suivantes à l'aide de la classification de configuration `presto-config`, qui correspond au fichier `config.properties` pour Presto :
  - Définit `http-server.http.enabled` sur `false` sur tous les nœuds, ce qui désactive HTTP en faveur de HTTPS. Cela nécessite que vous fournissiez des certificats qui fonctionnent pour le DNS public et privé lors de la configuration de sécurité pour le chiffrement en transit. Pour ce faire, vous pouvez utiliser des certificats SAN (Subject Alternative Name) qui prennent en charge plusieurs domaines.
  - Il définit les valeurs `http-server.https.*`. Pour plus de détails sur la configuration, consultez Authentification LDAP dans la documentation Presto.

De plus, avec les versions 5.10.0 et ultérieures d'Amazon EMR, vous pouvez configurer l'authentification LDAP pour les connexions client au coordinateur Presto à l'aide du protocole HTTPS. Cette configuration utilise LDAP sécurisé (LDAPS). TLS doit être activé sur votre serveur LDAP et le cluster Presto doit utiliser une configuration de sécurité avec le chiffrement des données en transit activé. Une configuration supplémentaire est requise. Les options de configuration sont différentes selon la version d'Amazon EMR que vous utilisez. Pour de plus amples informations, veuillez consulter Utilisation de l'authentification LDAP pour Presto sur Amazon EMR.

Presto sur Amazon EMR utilise par défaut le port 8446 pour le HTTPS interne. Le port utilisé pour les communications internes doit être le même port que celui utilisé pour l'accès HTTPS client au coordinateur Presto. La propriété `http-server.https.port` dans la classification de configuration `presto-config` spécifie le port.

Utilisation de l'authentification LDAP pour Presto sur Amazon EMR

Suivez les étapes de cette section pour configurer LDAP. Chaque étape contient des exemples et des liens vers des informations complémentaires.

Étapes de configuration de l'authentification LDAP

- Étape 1 : Collecter des informations sur votre serveur LDAP et copier le certificat du serveur sur Amazon S3
- Étape 2 : Définir une configuration de sécurité
- Étape 3 : Créer une configuration JSON avec des propriétés LDAP pour Presto

- Étape 4 : Créer le script pour copier le certificat de serveur LDAP et le charger dans Amazon S3
- Étape 5 : Créer le cluster

Étape 1 : Collecter des informations sur votre serveur LDAP et copier le certificat du serveur sur Amazon S3

Dans la section suivante, vous aurez besoin des informations et des éléments suivants relatifs à votre serveur LDAP pour configurer l'authentification LDAP.

L'adresse IP ou le nom d'hôte du serveur LDAP

Le coordinateur Presto sur le nœud principal Amazon EMR doit être en mesure d'atteindre le serveur LDAP via l'adresse IP ou le nom d'hôte spécifié. Par défaut, Presto communique avec le serveur LDAP via LDAPS sur le port 636. Si votre implémentation LDAP nécessite un port personnalisé, vous pouvez le spécifier à l'aide de la propriété `ldap.url` avec les versions 5.16.0 ou ultérieures d'Amazon EMR, ou en utilisant `authentication.ldap.url` avec les versions antérieures. Remplacez `636` par le port personnalisé comme illustré dans les exemples de configuration `presto-config` de Étape 3 : Créer une configuration JSON avec des propriétés LDAP pour Presto. Assurez-vous que les pare-feu et les groupes de sécurité autorisent le trafic entrant et sortant sur le port 636 (ou votre port personnalisé) ainsi que sur le port 8446 (ou votre port personnalisé), qui est utilisé pour les communications de cluster internes.

Le certificat de serveur LDAP

Vous devez charger le fichier de certificat dans un emplacement sécurisé dans Amazon S3. Pour plus d'informations, consultez Comment charger des fichiers ou des dossiers dans un compartiment S3 dans le Guide de l'utilisateur Amazon Simple Storage Service. Vous créez une action d'amorçage qui copie ce certificat depuis Amazon S3 vers chaque nœud du cluster lors du lancement du cluster. Dans Étape 4 : Créer le script pour copier le certificat de serveur LDAP et le charger dans Amazon S3. L'exemple de certificat est *`s3://amzn-s3-demo-bucket/ldap_server.crt`*.

Les paramètres du serveur LDAP pour des liaisons anonymes

Si la liaison anonyme est désactivée sur PrestoDB, vous avez besoin de l'ID utilisateur (UID) et du mot de passe d'un compte disposant des autorisations pour établir la liaison avec le serveur LDAP afin que le serveur PrestoDB puisse établir une connexion. Vous spécifiez l'UID et le mot de passe en utilisant les propriétés `internal-communication.authentication.ldap.user` et `internal-communication.authentication.ldap.password` dans la classification de configuration

`presto-config`. Amazon EMR 5.10.0 ne prend pas en charge ces paramètres ; la liaison anonyme doit donc être prise en charge sur le serveur LDAP lorsque vous utilisez cette version.

Notez que Trino ne nécessite pas la configuration de liaison anonyme.

Pour obtenir le statut de liaison anonyme sur le serveur LDAP

- Utilisez la commande ldapwhoami à partir d'un client Linux, comme illustré dans l'exemple suivant :

  ```
  ldapwhoami -x -H ldaps://LDAPServerHostNameOrIPAddress
  ```

  Si la liaison anonyme n'est pas autorisée, la commande renvoie les informations suivantes :

  ```
  ldap_bind: Inappropriate authentication (48)
  additional info: anonymous bind disallowed
  ```

Pour vérifier qu'un compte dispose des autorisations pour un serveur LDAP qui utilise l'authentification simple

- Utilisez la commande ldapwhoami à partir d'un client Linux, comme illustré dans l'exemple suivant. L'exemple utilise un utilisateur fictif*presto*, stocké sur un serveur Open LDAP exécuté sur une EC2 instance portant le nom d'hôte fictif. *ip-xxx-xxx-xxx-xxx.ec2.internal* L'utilisateur est associé à l'unité organisationnelle (UO) *admins* et au mot de passe *123456* :

  ```
  ldapwhoami -x -w "123456" -D uid=presto,ou=admins,dc=ec2,dc=internal -H ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal
  ```

  Si le compte est valide et dispose des autorisations appropriées, la commande renvoie :

  ```
  dn:uid=presto,ou=admins,dc=ec2,dc=internal
  ```

Les exemples de configuration dans Étape 3 : Créer une configuration JSON avec des propriétés LDAP pour Presto incluent ce compte pour plus de clarté, à l'exception de l'exemple 5.10.0, dans lequel il n'est pas pris en charge. Si le serveur LDAP utilise une liaison anonyme, supprimez les paires nom/valeur `internal-communication.authentication.ldap.user` et `internal-communication.authentication.ldap.password`.

Le nom unique LDAP pour les utilisateurs Presto

Lorsque vous spécifiez la configuration LDAP pour Presto, vous spécifiez un modèle de liaison composé d'`${USER}`une unité organisationnelle (UO) et de composants de domaine supplémentaires (DCs). Presto remplace `${USER}` par l'UD utilisateur effectif (UID) de chaque utilisateur pendant l'authentification par mot de passe afin qu'il corresponde au nom unique spécifié par ce modèle de liaison. Vous avez besoin de OUs ceux auxquels appartiennent les utilisateurs éligibles et de leurs DCs. Par exemple, pour autoriser les utilisateurs de l'unité d'organisation `admins` ou du domaine `corp.example.com` à s'authentifier auprès de Presto, vous spécifiez `${USER},ou=admins,dc=corp,dc=example,dc=com` en tant que modèle de liaison utilisateur.

> **Note**
>
> Lorsque vous utilisez AWS CloudFormation, vous devez utiliser la fonction Fn : :Sub afin de la `${USER}` remplacer par l'ID utilisateur (UID) réel. Pour plus d'informations, consultez la rubrique Fn::Sub dans le Guide de l'utilisateur AWS CloudFormation .

Lorsque vous utilisez Amazon EMR 5.10.0, vous ne pouvez spécifier qu'un seul modèle de ce type. Avec Amazon EMR 5.11.0 ou version ultérieure, vous pouvez spécifier plusieurs modèles séparés par un signe deux points (:). Les utilisateurs qui tentent de s'authentifier auprès de Presto sont comparés au premier modèle, puis au deuxième, et ainsi de suite. Pour obtenir un exemple, consultez Étape 3 : Créer une configuration JSON avec des propriétés LDAP pour Presto.

Étape 2 : Définir une configuration de sécurité

Créez une configuration de sécurité avec le chiffrement en transit activé. Pour plus d'informations, consultez Création d'une configuration de sécurité dans le Guide de gestion d'Amazon EMR. Les artefacts de chiffrement que vous fournissez lorsque vous configurez le chiffrement en transit sont utilisés pour chiffrer les communications internes entre les nœuds Presto. Pour plus d'informations, consultez Mise à disposition des certificats de chiffrement des données en transit. Le certificat de serveur LDAP est utilisé pour authentifier les connexions client auprès du serveur Presto.

Étape 3 : Créer une configuration JSON avec des propriétés LDAP pour Presto

Vous utilisez la classification de configuration `presto-config` pour définir des propriétés Presto pour LDAP. Le format et le contenu de `presto-config` sont légèrement différents selon la version d'Amazon EMR et l'installation de Presto (PrestoDB ou Trino). Des exemples de configuration

sont fournis plus tard dans cette section. Pour de plus amples informations, veuillez consulter Configuration des applications.

Les étapes suivantes supposent que vous enregistrez les données JSON dans un fichier*MyPrestoConfig.json*. Si vous utilisez la console, chargez le fichier dans un emplacement sécurisé dans Amazon S3 afin de pouvoir y faire référence lorsque vous créez le cluster. Si vous utilisez le AWS CLI, vous pouvez référencer le fichier localement.

Example Amazon EMR 6.1.0 et versions ultérieures avec PrestoSQL (Trino)

L'exemple suivant utilise le nom d'hôte LDAP de Étape 1 : Collecter des informations sur votre serveur LDAP et copier le certificat du serveur sur Amazon S3 pour s'authentifier auprès du serveur LDAP pour la liaison. Deux modèles de liaison utilisateur sont spécifiés, ce qui indique que les utilisateurs des unités d'organisation `admins` et `datascientists` sur le serveur LDAP sont éligibles pour l'authentification auprès du serveur Trino en tant qu'utilisateurs. Les modèles de liaison sont séparés par un signe deux-points (`:`).

Les versions 6.4.0 et ultérieures d'Amazon EMR utilisent le nouveau nom Trino au lieu de PrestoSQL. Si vous utilisez Trino, remplacez *`prestosql-config`* dans la classification de configuration suivante par `trino-config` et *`prestosql-password-authenticator`* par `trino-password-authenticator`.

```
[
   {
      "Classification":"prestosql-config",
      "Properties":{
         "http-server.authentication.type":"PASSWORD"
      }
   },
   {
      "Classification":"prestosql-password-authenticator",
      "Properties":{
         "password-authenticator.name":"ldap",
         "ldap.url":"ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal:636",
         "ldap.user-bind-pattern": "uid=${USER},ou=admins,dc=ec2,dc=internal:uid=
${USER},ou=datascientists,dc=ec2,dc=internal"
      }
   }
]
```

Example Amazon EMR 5.16.0 et versions ultérieures

L'exemple suivant utilise l'ID utilisateur et le mot de passe LDAP, ainsi que le nom d'hôte LDAP Étape 1 : Collecter des informations sur votre serveur LDAP et copier le certificat du serveur sur Amazon S3 pour l'authentification auprès du serveur LDAP pour la liaison. Deux modèles de liaison utilisateur sont spécifiés, ce qui indique que les utilisateurs des unités d'organisation `admins` et `datascientists` sur le serveur LDAP sont éligibles pour l'authentification auprès du serveur Presto en tant qu'utilisateurs. Les modèles de liaison sont séparés par un signe deux-points (`:`).

```
[{
        "Classification": "presto-config",
                "Properties": {
                        "http-server.authentication.type": "PASSWORD"
                }
        },
        {
                "Classification": "presto-password-authenticator",
                "Properties": {
                        "password-authenticator.name": "ldap",
                        "ldap.url": "ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal:636",
                        "ldap.user-bind-pattern": "uid=
${USER},ou=admins,dc=ec2,dc=internal:uid=${USER},ou=datascientists,dc=ec2,dc=internal",
                        "internal-communication.authentication.ldap.user": "presto",
                        "internal-communication.authentication.ldap.password": "123456"
                }
        }]
```

Example Amazon EMR version 5.11.0 à 5.15.0

Le format de la classification de configuration `presto-config` est légèrement différent pour ces versions. L'exemple suivant spécifie les mêmes paramètres que l'exemple précédent.

```
[{
        "Classification": "presto-config",
                "Properties": {
                        "http-server.authentication.type": "LDAP",
                        "authentication.ldap.url": "ldaps://ip-xxx-xxx-xxx-
xxx.ec2.internal:636",
                        "authentication.ldap.user-bind-pattern": "uid=
${USER},ou=admins,dc=ec2,dc=internal:uid=${USER},ou=datascientists,dc=ec2,dc=internal",
                        "internal-communication.authentication.ldap.user": "presto",
                        "internal-communication.authentication.ldap.password": "123456"
```

```
                }
        }]
```

Example Amazon EMR 5.10.0

Amazon EMR 5.10.0 prend en charge la liaison anonyme uniquement ; ces entrées sont donc omises. De plus, un seul modèle de liaison peut être spécifié.

```
[{
        "Classification": "presto-config",
                "Properties": {
                        "http-server.authentication.type": "LDAP",
                        "authentication.ldap.url": "ldaps://ip-xxx-xxx-xxx-
xxx.ec2.internal:636",
                        "ldap.user-bind-pattern": "uid=
${USER},ou=prestousers,dc=ec2,dc=internal"
                }
        }]
```

Étape 4 : Créer le script pour copier le certificat de serveur LDAP et le charger dans Amazon S3

Créez un script qui copie le fichier de certificat sur chaque nœud du cluster et l'ajoute au keystore. Créez le script à l'aide d'un éditeur de texte, enregistrez-le, puis chargez-le dans Amazon S3. DansÉtape 5 : Créer le cluster, le fichier de script est référencé sous la forme*s3://amzn-s3-demo-bucket/LoadLDAPCert.sh*.

L'exemple de script suivant utilise le mot de passe du keystore par défaut,*changeit*. Nous vous recommandons de vous connecter au nœud maître après avoir créé le cluster et de modifier le mot de passe du keystore à l'aide de la commande keytool.

```
#!/bin/bash
aws s3 cp s3://amzn-s3-demo-bucket/ldap_server.crt .
sudo keytool -import -keystore /usr/lib/jvm/jre-1.8.0-openjdk.x86_64/lib/security/
cacerts -trustcacerts -alias ldap_server -file ./ldap_server.crt -storepass changeit -
noprompt
```

Étape 5 : Créer le cluster

Lorsque vous créez le cluster, vous spécifiez Presto et les autres applications que vous voulez qu'Amazon EMR installe. Les exemples suivants font également référence aux propriétés de classification de configuration dans un fichier au format JSON, mais vous pouvez également spécifier la classification de configuration en ligne.

Pour créer un cluster Presto avec authentification LDAP à l'aide de la console Amazon EMR.

1. Accédez à la nouvelle console Amazon EMR et sélectionnez Changer pour l'ancienne console depuis le menu latéral. Pour plus d'informations sur ce qu'implique le passage à l'ancienne console, consultez la rubrique Utilisation de l'ancienne console.
2. Choisissez Créer un cluster et Go to advanced options (Aller aux options avancées).
3. Choisissez Presto et les autres applications devant être installées par Amazon EMR, puis sous Configuration logicielle, sélectionnez la version d'Amazon EMR à utiliser. L'authentification LDAP est prise en charge uniquement avec Amazon EMR version 5.10.0 et ultérieure.
4. Sous Modifier les paramètres logiciels, choisissez Charger JSON à partir de S3 , indiquez l'emplacement dans Amazon S3 du fichier de configuration JSON que vous avez créé dans Étape 3 : Créer une configuration JSON avec des propriétés LDAP pour Presto, puis choisissez Suivant.
5. Configurez le matériel et la mise en réseau du cluster, puis choisissez Next (Suivant).
6. Choisissez Bootstrap Actions (Actions d'amorçage). Pour Add bootstrap action (Ajouter une action d'amorçage), sélectionnez Custom action (Action personnalisée), puis choisissez Configure and add (Configurer et ajouter).
7. Entrez un nom pour l'action bootstrap, entrez l'emplacement du script dans lequel vous l'avez crééÉtape 4 : Créer le script pour copier le certificat de serveur LDAP et le charger dans Amazon S3, par exemple s3://amzn-s3-demo-bucket/Load LDAPCert .sh, puis choisissez Ajouter.
8. Sous General Options (Options générales), Tags (Balises) et Additional Options (Options supplémentaires), choisissez les paramètres appropriés à votre application, puis choisissez Next (Suivant).
9. Choisissez Authentication and encryption (Authentification et chiffrement), puis sélectionnez la configuration de sécurité que vous avez créée dans Étape 2 : Définir une configuration de sécurité.
10. Choisissez d'autres options de sécurité en fonction de votre application, puis choisissez Create cluster (Créer le cluster).

Pour créer un cluster Presto avec l'authentification LDAP à l'aide de l' AWS CLI

- Utilisez la commande `aws emr create-cluster`. Au minimum, spécifiez l'application Presto, ainsi que la classification de configuration Presto, le script d'amorçage et la configuration de sécurité que vous avez créée au cours des étapes précédentes. L'exemple suivant fait référence au fichier de configuration en tant que fichier JSON enregistré dans le même répertoire que celui où vous exécutez la commande. Le script d'amorçage, d'autre part, doit être enregistré dans Amazon S3. L'exemple suivant utilise `s3://amzn-s3-demo-bucket/LoadLDAPCert.sh`.

  > **Note**
  >
  > Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster  --applications Name=presto --release-label emr-5.16.0 \
--use-default-roles --ec2-attributes KeyName=MyKeyPair,SubnetId=subnet-1234ab5
 \ --instance-count 3 --instance-type m5.xlarge --region us-west-2 --name
 "MyPrestoWithLDAPAuth" \
--bootstrap-actions Name="Distribute LDAP server cert",Path="s3://amzn-s3-demo-
bucket/LoadLDAPCert.sh" \
--security-configuration MyPrestoLDAPSecCfg --configurations file://
MyPrestoConfig.json
```

Activation du mode strict Presto

Dans certaines situations, les requêtes de longue durée peuvent entraîner des coûts élevés et obliger Amazon EMR à utiliser davantage de ressources de cluster. Cela réduit les ressources des autres charges de travail du cluster. Avec les versions 6.8 et ultérieures d'Amazon EMR, vous pouvez utiliser une fonctionnalité de mode strict qui rejette ou vous avertit des types de requêtes de longue durée suivants :

- Requêtes sans prédicats sur les colonnes partitionnées qui entraînent l'analyse de grandes quantités de données dans les tables
- Requêtes avec jointures croisées entre deux grandes tables
- Requêtes qui trient un grand nombre de lignes sans limite

Une fois que Presto a complètement optimisé le plan de requêtes, le mode strict s'exécute. Pour utiliser et personnaliser le mode strict en fonction des besoins de vos requêtes, vous pouvez configurer Presto de la manière suivante.

Configurations Presto pour le mode strict

Paramètre	Description	Par défaut
`strict-mode-enabled`	Active et désactive le mode strict. La valeur de `true` indique que le mode strict est activé.	`false`
`strict-mode-fail-query`	Rejette les requêtes si le mode strict détecte des requêtes probables de longue durée. Si `false`, Amazon EMR émet uniquement un avertissement.	`false`
`strict-mode-restrictions`	Spécifie les restrictions à appliquer lorsque le mode strict est activé. Le mode strict prend en charge les restrictions suivantes : MANDATORY_PARTITION_PREDICATE, DISALLOW_CROSS_JOIN et LIMITED_SORT.	MANDATORY_PARTITION_PREDICATE, DISALLOW_CROSS_JOIN, LIMITED_SORT

Pour tester le mode strict, vous pouvez remplacer ces configurations ou les définir comme propriétés de session lorsque vous utilisez le client Presto.

Pour définir la configuration lors de la création du cluster à l'aide de la AWS Management Console

1. Choisissez Créer un cluster et sélectionnez Amazon EMR version 6.8.0, puis Presto ou Trino. Pour plus d'informations, veuillez consulter Installation de PrestoDB et Trino.
2. Spécifiez directement les propriétés de configuration pour le mode strict ou chargez un fichier JSON sur Amazon S3. Vous pouvez éventuellement sélectionner le catalogue de données AWS

Glue pour votre métastore. Spécifiez votre VPC, vos sous-réseaux, vos actions d'amorçage, votre paire de clés et votre groupe de sécurité. Choisissez Créer un cluster pour créer votre cluster.

3. Connectez-vous au nœud primaire du cluster et exécutez `presto-cli` ou `trino-cli`.
4. Soumettez vos questions. Le mode strict valide chaque requête et détermine si elle est de longue durée. En fonction de vos paramètres `strict-mode-fail-query`, Amazon EMR rejette la requête ou émet un avertissement.
5. Lorsque vous avez terminé vos requêtes, résiliez le cluster et supprimez vos ressources.

Pour définir la configuration sur un cluster en cours d'exécution à l'aide de la AWS CLI

1. Connectez-vous au nœud principal de votre cluster à l'aide du AWS CLI et exécutez `presto-cli` ou`trino-cli`.
2. Exécutez les commandes suivantes avec les valeurs souhaitées.

```
set session strict_mode_enabled = true;
set session strict_mode_fail_query = false;
set session strict_mode_restrictions = 'DISALLOW_CROSS_JOIN,LIMITED_SORT';
```

Considérations

Lorsque vous utilisez le mode strict, tenez compte des éléments suivants :

- Dans certains cas, le mode strict peut rejeter les requêtes de courte durée qui ne consomment pas beaucoup de ressources. Par exemple, les requêtes sur de petites tables n'appliquent pas de filtrage dynamique et ne remplacent pas les jointures internes par des jointures croisées. Cela peut amener la requête à utiliser le prédicat de partition obligatoire ou à interdire la jointure croisée. Dans ce cas, le mode strict rejette la requête.
- Le contrôle du mode strict s'applique uniquement aux types de requêtes SELECT, INSERT, CREATE TABLE AS SELECT et EXPLAIN ANALYZE.
- Vous ne pouvez utiliser la restriction sur le prédicat de partition obligatoire qu'avec le connecteur Hive.

Gestion de la perte d'instances Spot dans Presto

Avec les instances Spot d'Amazon EMR, vous pouvez exécuter des charges de travail de big data sur la EC2 capacité inutilisée d'Amazon à moindre coût. En échange de cette réduction des coûts, Amazon EC2 peut interrompre les instances Spot avec une notification de deux minutes. Lorsque vous résiliez un nœud, Presto peut prendre jusqu'à 10 minutes avant qu'il ne renvoie une erreur. Cela entraîne des retards inutiles dans les rapports d'erreur et les éventuelles nouvelles tentatives. La résiliation rapide est une fonctionnalité qui vous permet de contrôler la façon dont Presto gère les nœuds résiliés.

Le rôle du coordinateur Presto est de suivre tous les composants master en interrogeant régulièrement leur statut. Sans résiliation, le coordinateur ne consulte pas le YARN NodeManager pour connaître l'état de chaque nœud. Cela peut entraîner une longue boucle de nouvelles tentatives avant l'échec de la requête. En cas de résiliation rapide, le coordinateur Presto consulte l'état du nœud NodeManager dès que le sondage ne parvient pas à atteindre l'hôte. S' NodeManager il s'avère que le nœud est inactif, Presto abandonne toute nouvelle tentative, échoue à la requête et renvoie une erreur. `NODE_DECOMMISSIONED`

L'ensemble de paramètres de configuration suivant vous permet de contrôler et de personnaliser le comportement de Presto en cas de résiliation du nœud.

Configurations Presto pour la gestion des défaillances des nœuds

Paramètre	Description	Par défaut
`query.remote-task.max-backoff-duration`	Durée pendant laquelle le coordinateur continue de tenter de récupérer l'état de la tâche distante à partir des composants master.	10 minutes
`query.remote-task.quick-terminate-node-failure`	Active la défaillance rapide du nœud si le coordinateur ne peut pas atteindre le nœud ou ne peut pas se connecter au travailleur qui s'exécute sur ce nœud. La valeur de `query.remote-task.terminate-on-conne`	`true`

Paramètre	Description	Par défaut
	`ct-exception` détermine si le coordinateur doit atteindre le nœud ou se connecter au travailleur. Le nœud échoue à la requête et Amazon EMR le supprime de la liste des travailleurs disponibles. Dans ce cas, vous ne pouvez pas utiliser le nœud pour planifier de nouvelles requêtes. Lorsque vous définissez cette valeur sur `false`, Presto revient à son comportement précédent où le coordinat eur Presto essaie à nouveau d'atteindre le nœud (pour `query.remote-task. max-backoff-durati on`) avant de marquer le nœud comme indisponible et d'échouer à la requête en cours sur le nœud.	

Paramètre	Description	Par défaut
`query.remote-task.terminate-on-connect-exception`	Spécifie si Amazon EMR doit créer un nœud si l'hôte est accessible mais que le coordinateur ne parvient pas à se connecter au processus de travail de l'hôte. Lorsque vous définissez cette valeur sur `true`, vous activez l'échec rapide de la requête si l'hôte est inaccessible.	`false`

Utilisation du dimensionnement automatique de Presto avec désaffectation gracieuse

Les versions 5.30.0 et ultérieures d'Amazon EMR incluent une fonctionnalité que vous pouvez utiliser pour définir une période de grâce pour certaines actions de dimensionnement. La période de grâce permet aux tâches Presto de continuer à s'exécuter avant que le nœud ne se résilie en raison d'une action de redimensionnement ou d'une demande de politique de dimensionnement automatique. Pour plus d'informations sur les règles de dimensionnement, consultez Comprendre les règles de dimensionnement automatique dans le Guide de gestion Amazon EMR. Presto autoscaling avec la mise hors service gracieuse empêche la planification de nouvelles tâches sur un nœud en cours de mise hors service, tout en permettant aux tâches déjà en cours d'exécution de se terminer avant que le délai d'arrêt ne soit atteint. Les requêtes en cours d'exécution se termineront avant que le nœud soit désaffecté. Autoscaling n'est pas pris en charge sur les parcs d'instances.

Vous pouvez contrôler le délai d'exécution des tâches Presto après réception d'une demande d'arrêt automatique. Par défaut, le délai d'arrêt d'Amazon EMR est de `0` minutes, ce qui signifie qu'Amazon EMR résilie immédiatement le nœud et toutes les tâches Presto en cours d'exécution sur celui-ci, si une demande de mise à l'échelle l'exige. Pour définir un délai plus long pour les tâches Presto sur Amazon EMR afin de permettre aux requêtes en cours d'exécution de se terminer avant la réduction de capacité d'un cluster, utilisez la classification de configuration `presto-config` pour définir le paramètre `graceful-shutdown-timeout` sur une valeur (en secondes ou en minutes) supérieure à zéro. Pour de plus amples informations, veuillez consulter Configuration des applications.

Par exemple, l'augmentation de la valeur `graceful-shutdown-timeout` à `"30m"` spécifie un délai d'attente de 30 minutes. À la fin du délai d'arrêt, le nœud marqué pour la désaffectation fait l'objet d'un arrêt forcé s'il attend que des tâches de requête se terminent. Dans ce cas, la requête échoue. Si les tâches de requête se terminent au bout de cinq minutes, le nœud marqué pour la mise hors service est résilié au bout de cinq minutes, à condition que les autres applications YARN aient terminé leur exécution.

Example Exemple de configuration de dimensionnement automatique Presto avec désaffectation gracieuse

Remplacez la valeur `graceful-shutdown-timeout` par le nombre de minutes approprié pour votre configuration. Il n'y a pas de valeur maximale. L'exemple ci-dessous définit une valeur de délai d'arrêt de `1800` secondes (30 minutes).

```
[
    {
        "classification": "presto-config",
        "properties": {
            "graceful-shutdown-timeout": "1800s"
        }
    }
]
```

Limites

La mise hors service gracieuse de PrestoDB ne fonctionne pas sur les clusters EMR où la connectivité HTTP est désactivée, par exemple quand `http-server.http.enabled` est défini sur `false`. Trino ne prend pas du tout en charge la mise hors service gracieuse, quel que soit le paramètre `http-server.http.enabled`.

Considérations relatives à Presto sur Amazon EMR

Tenez compte limitations suivantes lors de l'exécution de Presto sur Amazon EMR.

Exécutable en ligne de commande Presto

Dans Amazon EMR, PrestoDB et Trino utilisent tous deux le même exécutable en ligne de commande, `presto-cli`, comme dans l'exemple suivant.

```
presto-cli --catalog hive
```

Propriétés de déploiement Presto non configurables

La version d'Amazon EMR que vous utilisez détermine les configurations de déploiement Presto disponibles. Pour plus d'informations sur ces propriétés de configuration, consultez la section Déploiement de Presto dans la documentation de Presto. Le tableau suivant montre les différentes options de configuration pour les fichiers Presto `properties`.

Fichier	Configurable
`log.properties`	PrestoDB : configurable dans les versions 4.0.0 et ultérieures d'Amazon EMR. Utilisez la classification de configuration `presto-log` .
`config.properties`	PrestoDB : configurable dans les versions 4.0.0 et ultérieures d'Amazon EMR. Utilisez la classification de configuration `presto-config` .
`hive.properties`	PrestoDB : configurable dans les versions 4.1.0 et ultérieures d'Amazon EMR. Utilisez la classification de configuration `presto-connector-hive` .
`node.properties`	PrestoDB : configurable dans les versions 5.6.0 et ultérieures d'Amazon EMR. Utilisez la classification de configuration `presto-node` .
`jvm.config`	Non configurable.

Installation de Prestodb

Le nom de l'application Presto continue d'être utilisé pour installer PrestoDB sur des clusters.

Vous pouvez installer PrestoDB ou Trino, mais vous ne pouvez pas installer les deux sur un seul cluster. Si vous spécifiez à la fois Prestodb et Trino lorsque vous tentez de créer un cluster, une erreur de validation se produit et la demande de création de cluster échoue.

Configuration d'EMRFS et de PrestOS3 FileSystem

Avec les versions 5.12.0 et ultérieures d'Amazon EMR, PrestoDB peut utiliser EMRFS. Pour plus d'informations, consultez EMR File System (EMRFS) dans le Guide de gestion Amazon EMR. Dans les versions antérieures d'Amazon EMR, PrestOS3 FileSystem est la seule option de configuration.

Vous pouvez utiliser une configuration de sécurité pour configurer le chiffrement des données EMRFS dans Amazon S3. Vous pouvez également utiliser les rôles IAM pour les demandes EMRFS à Amazon S3. Pour plus d'informations, consultez les sections Comprendre les options de chiffrement et Configuration des rôles IAM pour les demandes EMRFS vers Amazon S3 dans le Guide de gestion Amazon EMR.

> **Note**
>
> Si vous interrogez des données sous-jacentes dans Amazon S3 avec Amazon EMR version 5.12.0, des erreurs Presto peuvent se produire. Cela est dû au fait que Presto ne parvient pas à récupérer les valeurs de classification de configuration depuis `emrfs-site.xml`. Pour contourner le problème, créez un sous-répertoire `emrfs` sous `usr/lib/presto/plugin/hive-hadoop2/` et un lien symbolique dans `usr/lib/presto/plugin/hive-hadoop2/emrfs` vers le fichier `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` existant. Redémarrez ensuite le processus presto-server (`sudo presto-server stop` suivi de `sudo presto-server start`).

Vous pouvez remplacer la valeur par défaut d'EMRFS et utiliser PrestOS3 à la place. FileSystem Pour ce faire, utilisez la classification de configuration `presto-connector-hive` pour définir `hive.s3-file-system-type` sur `PRESTO` comme illustré dans l'exemple suivant. Pour de plus amples informations, veuillez consulter Configuration des applications.

```
[
   {
      "Classification": "presto-connector-hive",
      "Properties": {
         "hive.s3-file-system-type": "PRESTO"
      }
   }
]
```

Si vous utilisez PrestOS3FileSystem, utilisez la classification de `presto-connector-hive` configuration pour configurer les propriétés de PrestOS3. FileSystem Pour plus d'informations sur les propriétés disponibles, consultez Configuration Amazon S3 dans la section du connecteur Hive de la documentation Presto. Ces paramètres ne s'appliquent pas à EMRFS.

Paramètre par défaut de l'emprunt d'identité de l'utilisateur final

Par défaut, les versions 5.12.0 et ultérieures d'Amazon EMR activent l'emprunt d'identité de l'utilisateur final pour l'accès à HDFS. Pour plus d'informations, consultez Emprunt de l'identité de l'utilisateur final dans la documentation Presto. Pour modifier ce paramètre avec la classification de configuration `presto-config`, définissez la propriété `hive.hdfs.impersonation.enabled` sur `false`.

Port par défaut de l'interface Web Presto

Par défaut, Amazon EMR configure l'interface web Presto sur le coordinateur Presto pour utiliser le port 8889 (pour PrestoDB et Trino). Pour modifier le port, utilisez la classification de configuration `presto-config` pour définir la propriété `http-server.http.port`. Pour plus d'informations, consultez Propriétés de configuration dans la section Déploiement de Presto de la documentation Presto.

Problème d'exécution du compartiment Hive dans certaines versions

Presto version 152.3 connaît un problème lié à l'exécution de compartiment Hive qui entraîne des performances des requêtes Presto significativement plus lentes dans certaines circonstances. Les versions 5.0.3, 5.1.0 et 5.2.0 d'Amazon EMR incluent cette version de Presto. Pour atténuer ce problème, utilisez la classification de configuration `presto-connector-hive` pour définir la propriété `hive.bucket-execution` sur `false` tel qu'illustré dans l'exemple suivant.

```
[
   {
      "Classification": "presto-connector-hive",
      "Properties": {
         "hive.bucket-execution": "false"
      }
   }
]
```

Historique des versions de Presto

Le tableau suivant répertorie la version de Presto incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version de Presto

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-7.9.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.8.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-7,7.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.6.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-7.5.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.4.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-7.3.0	0,285	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.2.0	0,285	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.36.2	0,267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.1.0	0,284	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-7.0.0	0,283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.15.0	0,283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.14.0	0,281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.13.0	0,281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.12.0	0,281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.11.1	0,279	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.11.0	0,279	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.10.1	0,278	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.10.0	0,278	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.9.1	0.276	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.9.0	0.276	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.8.1	0,273	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.8.0	0,273	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.7.0	0,272	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.36.1	0,267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.36.0	0,267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.6.0	0,267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.35.0	0,266	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.5.0	0,261	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.4.0	0.254.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.3.1	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.3.0	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.2.1	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.2.0	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.1.1	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.1.0	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-6.0.1	0.230	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.0.0	0.230	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.34.0	0,261	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.33.1	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.33.0	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.32.1	0.240.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.32.0	0.240.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.31.1	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.31.0	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.30.2	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.30.1	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.30.0	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
mr-5.29.0	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.28.1	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.28.0	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.27.1	0.224	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.27.0	0.224	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.26.0	0.220	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.25.0	0.220	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.24.1	0.219	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.24.0	0.219	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.23.1	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.23.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.22.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.21.2	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.21.1	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.21.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.20.1	0.214	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.20.0	0.214	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.19.1	0.212	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.19.0	0.212	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.18.1	0.210	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.18.0	0.210	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.17.2	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.17.1	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.17.0	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.16.1	0.203	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.16.0	0.203	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.15.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.15.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.14.2	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.14.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.14.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.13.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.13.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.12.3	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.12.2	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.12.1	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.12.0	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.4	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.11.3	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.2	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.11.1	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.0	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.10.1	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.10.0	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.9.1	0.184	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.9.0	0.184	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.8.3	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.8.2	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.8.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.8.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.7.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.7.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.6.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.6.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.4	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.5.3	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.2	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.5.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.4.1	0.166	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.4.0	0.166	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.3.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.3.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.3.0	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.2.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.2.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.2.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.2.0	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.1.1	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.1.0	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.0.3	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.0.2	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.0.1	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Étiquette de version Amazon EMR	Version Presto	Composants installés avec Presto
emr-5.0.0	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Apache Spark

Apache Spark est un modèle distribué de programmation et d'infrastructure qui vous permet d'effectuer des opérations de machine learning, de traitement de flux ou d'analyse graphique avec les clusters Amazon EMR. De manière analogue à Apache Hadoop, Spark est un système de traitement distribué open source, couramment utilisé pour les charges de travail de big data. Cependant, Spark présente plusieurs différences notables par rapport à Hadoop MapReduce. Spark a un moteur d'exécution optimisé de graphes acycliques dirigés (DAG) et met en cache de façon active les données en mémoire, ce qui peut améliorer les performances, notamment pour certains algorithmes et requêtes interactives.

De manière native, Spark prend en charge les applications écrites en Scala, python ou Java. Il inclut également plusieurs bibliothèques étroitement intégrées pour SQL (Spark), le machine learning (MLlib), le traitement de flux (Spark streaming) et le traitement de graphes (GraphX). Ces outils facilitent l'exploitation de l'infrastructure Spark pour une grande variété de cas d'utilisation.

Vous pouvez installer Spark sur un cluster Amazon EMR avec d'autres applications Hadoop. Cette solution peut également exploiter le système de fichiers Amazon EMR (EMRFS) pour accéder directement aux données dans Amazon S3. Hive est également intégré à Spark afin que vous puissiez utiliser un HiveContext objet pour exécuter des scripts Hive à l'aide de Spark. Un contexte Hive est inclus dans le shell de Spark en tant que `sqlContext`.

Pour un exemple de didacticiel sur la configuration d'un cluster EMR avec Spark et sur l'analyse d'un exemple de jeu de données, consultez Tutoriel : Getting started with Amazon EMR sur le blog News. AWS

> **Important**
>
> Apache Spark version 2.3.1, disponible à partir de la version 5.16.0 d'Amazon EMR, corrige CVE-2018-8024 et CVE-2018-1334. Nous vous recommandons de migrer les versions antérieures de Spark vers la version 2.3.1 ou ultérieure.

Le tableau suivant répertorie la version de Spark incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants qu'Amazon EMR installe avec Spark.

Pour connaître la version des composants installés avec Spark dans cette version, consultez la version 7.9.0 Versions des composants.

Informations sur la version de Spark pour emr-7.9.0

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-7.9.0	Spark 3.5.5	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Le tableau suivant répertorie la version de Spark incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Spark.

Pour connaître la version des composants installés avec Spark dans cette version, consultez la rubrique Versions des composants de la version 6.15.0.

Informations sur la version de Spark pour emr-6.15.0

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.15.0	Spark 3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
		kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

> **Note**
>
> La version 6.8.0 d'Amazon EMR est fournie avec la version 3.3.0 d'Apache Spark. Cette version de Spark utilise Apache Log4j 2 et le fichier `log4j2.properties` pour configurer Log4j dans les processus Spark. Si vous utilisez Spark dans le cluster ou si vous créez des clusters EMR avec des paramètres de configuration personnalisés, et que vous voulez passer à la version 6.8.0 d'Amazon EMR, vous devez migrer vers la nouvelle classification de configuration `spark-log4j2` et le nouveau format de clé pour Apache Log4j 2. Pour de plus amples informations, veuillez consulter Migration d'Apache Log4j 1.x vers Log4j 2.x.

Le tableau suivant répertorie la version de Spark incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Spark.

Pour la version des composants installés avec Spark dans cette version, consultez la version 5.36.2 Versions des composants.

Informations sur la version de Spark pour emr-5.36.2

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.36.2	Spark 2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode,

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
		hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Rubriques

- Création d’un cluster avec Apache Spark
- Exécution d’applications Spark avec Docker sur Amazon EMR 6.x
- Utiliser le catalogue AWS Glue Data Catalog avec Spark sur Amazon EMR
- Utilisation d'une hiérarchie de plusieurs catalogues dans AWS Glue Data Catalog avec Spark sur Amazon EMR
- Configuration de Spark
- Optimisation des performances de Spark
- Mise en cache des fragments de résultats Spark
- Utilisation de l’accélérateur Nvidia RAPIDS pour Apache Spark
- Accès au shell de Spark
- Utiliser Amazon SageMaker Spark pour le machine learning
- Ecriture d'une application Spark
- Améliorer les performances de Spark avec Amazon S3
- Ajout d'une étape Spark
- Afficher l'historique de l'application Spark
- Accédez au site Web de Spark UIs

- Utilisation du connecteur Amazon Kinesis Data Streams pour le streaming structuré Spark
- Utilisation de l'intégration d'Amazon Redshift pour Apache Spark avec Amazon EMR
- Historique des versions de Spark

Création d’un cluster avec Apache Spark

La procédure suivante permet de créer un cluster avec Spark installé à l’aide des options rapides de la console Amazon EMR.

Vous pouvez alternativement utiliser Options avancées pour personnaliser davantage la configuration de votre cluster, ou pour soumettre des étapes d'installation par programmation d'applications, puis d'exécution d'applications personnalisées. Avec l'une ou l'autre des options de création de cluster, vous pouvez choisir d'utiliser AWS Glue comme métastore Spark SQL. Pour plus d’informations, consultez Utiliser le catalogue AWS Glue Data Catalog avec Spark sur Amazon EMR.

Pour lancer un cluster avec Spark installé

1. Ouvrez la console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com
2. Choisissez Créer un cluster pour utiliser Options rapides.
3. Entrez un nom de cluster. Le nom de votre cluster ne peut pas contenir les caractères <, >, $, | ou `(backtick).
4. Pour Configuration logicielle, choisissez une option de Version.
5. Sous Applications, choisissez la solution groupée d'applications Spark.
6. Sélectionnez d'autres options si nécessaire, puis choisissez Create cluster (Créer le cluster).

> **Note**
>
> Pour configurer Spark lorsque vous créez le cluster, consultez Configuration de Spark.

Pour lancer un cluster sur lequel Spark est installé à l'aide du AWS CLI

- Créez le cluster à l'aide de la commande suivante.

```
aws emr create-cluster --name "Spark cluster" --release-label emr-7.9.0 --
applications Name=Spark \
```

```
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 --use-
default-roles
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

Pour lancer un cluster avec Spark installé à l'aide du kit SDK pour Java

Spécifiez Spark comme application avec `SupportedProductConfig` utilisé dans `RunJobFlowRequest`.

- L'exemple suivant montre comment créer un cluster avec Spark au moyen de Java.

```
import com.amazonaws.AmazonClientException;
import com.amazonaws.auth.AWSCredentials;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.profile.ProfileCredentialsProvider;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduce;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduceClientBuilder;
import com.amazonaws.services.elasticmapreduce.model.*;
import com.amazonaws.services.elasticmapreduce.util.StepFactory;

public class Main {

        public static void main(String[] args) {
                AWSCredentials credentials_profile = null;
                try {
                        credentials_profile = new
 ProfileCredentialsProvider("default").getCredentials();
                } catch (Exception e) {
                        throw new AmazonClientException(
                                        "Cannot load credentials from .aws/
credentials file. " +
                                                        "Make sure that the
 credentials file exists and the profile name is specified within it.",
                                        e);
```

```
                }

                AmazonElasticMapReduce emr =
 AmazonElasticMapReduceClientBuilder.standard()
                                .withCredentials(new
 AWSStaticCredentialsProvider(credentials_profile))
                                .withRegion(Regions.US_WEST_1)
                                .build();

                // create a step to enable debugging in the AWS Management Console
                StepFactory stepFactory = new StepFactory();
                StepConfig enabledebugging = new StepConfig()
                                .withName("Enable debugging")
                                .withActionOnFailure("TERMINATE_JOB_FLOW")

 .withHadoopJarStep(stepFactory.newEnableDebuggingStep());

                Application spark = new Application().withName("Spark");

                RunJobFlowRequest request = new RunJobFlowRequest()
                                .withName("Spark Cluster")
                                .withReleaseLabel("emr-5.20.0")
                                .withSteps(enabledebugging)
                                .withApplications(spark)
                                .withLogUri("s3://path/to/my/logs/")
                                .withServiceRole("EMR_DefaultRole")
                                .withJobFlowRole("EMR_EC2_DefaultRole")
                                .withInstances(new JobFlowInstancesConfig()
                                                .withEc2SubnetId("subnet-12ab3c45")
                                                .withEc2KeyName("myEc2Key")
                                                .withInstanceCount(3)

 .withKeepJobFlowAliveWhenNoSteps(true)
                                                .withMasterInstanceType("m4.large")

 .withSlaveInstanceType("m4.large"));
                RunJobFlowResult result = emr.runJobFlow(request);
                System.out.println("The cluster ID is " + result.toString());
        }
}
```

Exécution d'applications Spark avec Docker sur Amazon EMR 6.x

Avec Amazon EMR 6.0.0, les applications Spark peuvent utiliser des conteneurs Docker pour définir les dépendances de leurs bibliothèques, au lieu d'installer des dépendances sur les EC2 instances Amazon individuelles du cluster. Pour exécuter Spark avec Docker, vous devez d'abord configurer le registre Docker et définir des paramètres supplémentaires lors de la soumission d'une application Spark. Pour plus d'informations, consultez Configuration de l'intégration Docker.

Lorsque l'application est soumise, YARN appelle Docker pour extraire l'image Docker spécifiée et exécuter l'application Spark dans un conteneur Docker. Cette approche vous permet de définir et d'isoler facilement les dépendances. Elle réduit le temps d'amorçage ou de préparation des instances dans le cluster Amazon EMR avec les bibliothèques nécessaires à l'exécution des tâches.

Considérations lors de l'exécution de Spark avec Docker

Lorsque vous exécutez Spark avec Docker, assurez-vous que les conditions suivantes sont remplies :

- Le package `docker` et l'interface de ligne de commande ne sont installés que sur les nœuds principaux et de tâche.
- Sur Amazon EMR 6.1.0 et versions ultérieures, vous pouvez également installer Docker sur un nœud primaire à l'aide des commandes suivantes.
  - ```
    sudo yum install -y docker
    sudo systemctl start docker
    ```
- La commande `spark-submit` doit toujours être exécutée à partir d'une instance principale sur le cluster Amazon EMR.
- Les registres Docker utilisés pour résoudre les images Docker doivent être définis à l'aide de l'API Classification avec la clé de classification `container-executor` pour définir des paramètres supplémentaires lors du lancement du cluster :
  - `docker.trusted.registries`
  - `docker.privileged-containers.registries`
- Pour exécuter une application Spark dans un conteneur Docker, les options de configuration suivantes sont nécessaires :
  - `YARN_CONTAINER_RUNTIME_TYPE=docker`
  - `YARN_CONTAINER_RUNTIME_DOCKER_IMAGE={DOCKER_IMAGE_NAME}`
    ```

- Lorsque vous utilisez Amazon ECR pour récupérer des images Docker, vous devez configurer le cluster pour qu'il s'authentifie. Pour ce faire, vous devez utiliser l'option de configuration suivante :
  - YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG={DOCKER_CLIENT_CONFIG_PATH_ON
- Dans Amazon EMR 6.1.0 et versions ultérieures, vous n'êtes pas obligé d'utiliser la commande répertoriée `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG={DOCKER_CLIENT_CONFIG_PATH_ON_HDFS` lorsque la fonctionnalité d'authentification automatique ECR est activée.
- Toute image Docker utilisée avec Spark implique l'installation préalable de Java dans l'image Docker.

Pour plus d'informations sur les conditions préalables, consultez Configuration de l'intégration de Docker.

Création d'une image Docker

Les images Docker sont créées à l'aide d'un Dockerfile, qui définit les packages et la configuration à inclure dans l'image. Les deux exemples suivants utilisent Dockerfiles et PySpark SparkR.

PySpark Dockerfile

Les images Docker créées à partir de ce Dockerfile incluent Python 3 et le package Python NumPy . Ce Dockerfile utilise Amazon Linux 2 et le kit JDK Amazon Corretto 8.

```
FROM amazoncorretto:8

RUN yum -y update
RUN yum -y install yum-utils
RUN yum -y groupinstall development

RUN yum list python3*
RUN yum -y install python3 python3-dev python3-pip python3-virtualenv

RUN python -V
RUN python3 -V

ENV PYSPARK_DRIVER_PYTHON python3
ENV PYSPARK_PYTHON python3

RUN pip3 install --upgrade pip
```

```
RUN pip3 install numpy pandas

RUN python3 -c "import numpy as np"
```

Dockerfile SparkR

Les images Docker créées à partir de ce Dockerfile incluent R et le package RandomForest CRAN. Ce Dockerfile comprend Amazon Linux 2 et le kit JDK Amazon Corretto 8.

```
FROM amazoncorretto:8

RUN java -version

RUN yum -y update
RUN amazon-linux-extras install R4

RUN yum -y install curl hostname

#setup R configs
RUN echo "r <- getOption('repos'); r['CRAN'] <- 'http://cran.us.r-project.org';
 options(repos = r);" > ~/.Rprofile

RUN Rscript -e "install.packages('randomForest')"
```

Pour plus d'informations sur la syntaxe Dockerfile, consultez la documentation de référence relative à Dockerfile.

Utilisation d'images Docker depuis Amazon ECR

Amazon Elastic Container Registry (Amazon ECR) est un registre de conteneurs Docker entièrement géré qui permet de stocker, de gérer et de déployer facilement les images des conteneurs Docker. Lors de l'utilisation d'Amazon ECR, le cluster doit être configuré pour faire confiance à votre instance d'ECR. Vous devez, par ailleurs, configurer l'authentification pour que le cluster utilise les images Docker à partir d'Amazon ECR. Pour plus d'informations, consultez Configuration de YARN pour accéder à Amazon ECR.

Pour vous assurer que les hôtes Amazon EMR ont accès aux images stockées dans Amazon ECR, vérifiez que votre cluster dispose des autorisations de la politique `AmazonEC2ContainerRegistryReadOnly` associée au profil d'instance. Pour de plus amples informations, veuillez consulter Stratégie `AmazonEC2ContainerRegistryReadOnly`.

Dans cet exemple, le cluster doit être créé avec la configuration supplémentaire suivante pour s'assurer que le registre Amazon ECR est approuvé. Remplacez le *123456789123.dkr.ecr.us-east-1.amazonaws.com* point de terminaison par votre point de terminaison Amazon ECR.

```
[
  {
    "Classification": "container-executor",
    "Configurations": [
      {
        "Classification": "docker",
        "Properties": {
          "docker.privileged-containers.registries":
 "local,centos,123456789123.dkr.ecr.us-east-1.amazonaws.com",
          "docker.trusted.registries": "local,centos,123456789123.dkr.ecr.us-
east-1.amazonaws.com"
        }
      }
    ],
    "Properties": {}
  }
]
```

Utilisation PySpark avec Amazon ECR

L'exemple suivant utilise le PySpark Dockerfile, qui sera balisé et chargé sur Amazon ECR. Après avoir chargé le Dockerfile, vous pouvez exécuter la PySpark tâche et vous référer à l'image Docker depuis Amazon ECR.

Après avoir lancé le cluster, utilisez SSH pour vous connecter à un nœud principal et exécutez les commandes suivantes pour créer l'image Docker locale à partir de l'exemple PySpark Dockerfile.

Tout d'abord, créez un répertoire et un fichier Dockerfile.

```
mkdir pyspark
vi pyspark/Dockerfile
```

Collez le contenu du PySpark Dockerfile et exécutez les commandes suivantes pour créer une image Docker.

```
sudo docker build -t local/pyspark-example pyspark/
```

Créez le référentiel ECR `emr-docker-examples` pour les exemples.

```
aws ecr create-repository --repository-name emr-docker-examples
```

Marquez et téléchargez l'image créée localement sur ECR, en la remplaçant par votre point de *123456789123.dkr.ecr.us-east-1.amazonaws.com* terminaison ECR.

```
sudo docker tag local/pyspark-example 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-
docker-examples:pyspark-example
sudo docker push 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:pyspark-example
```

Utilisez SSH pour vous connecter au nœud primaire et préparez un script Python dont le nom de fichier est `main.py`. Collez le contenu suivant dans le fichier `main.py` et enregistrez-le.

```
from pyspark.sql import SparkSession
spark = SparkSession.builder.appName("docker-numpy").getOrCreate()
sc = spark.sparkContext

import numpy as np
a = np.arange(15).reshape(3, 5)
print(a)
```

Sur Amazon EMR 6.0.0, indiquez le nom de l'image Docker pour soumettre la tâche. Définissez les paramètres de configuration supplémentaires pour vous assurer que l'exécution de la tâche utilise Docker comme environnement d'exécution. Lors de l'utilisation d'Amazon ECR, `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG` doit référencer le fichier `config.json` contenant les informations d'identification utilisées pour s'authentifier auprès d'Amazon ECR.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:pyspark-example
DOCKER_CLIENT_CONFIG=hdfs:///user/hadoop/config.json
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
```

```
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
--num-executors 2 \
main.py -v
```

Sur Amazon EMR 6.1.0 et versions ultérieures, indiquez le nom de l'image Docker pour soumettre la tâche. Lorsque l'authentification automatique ECR est activée, veuillez exécuter la commande suivante.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:pyspark-example
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--num-executors 2 \
main.py -v
```

Lorsque la tâche est terminée, prenez note de l'ID de l'application YARN et utilisez la commande suivante pour obtenir le résultat de la PySpark tâche.

```
yarn logs --applicationId application_id | grep -C2 '\[\['
LogLength:55
LogContents:
[[ 0  1  2  3  4]
 [ 5  6  7  8  9]
 [10 11 12 13 14]]
```

Utilisation de SparkR avec Amazon ECR

L'exemple suivant utilise le fichier Dockerfile SparkR, qui sera balisé et chargé dans ECR. Une fois le fichier Dockerfile chargé, vous pouvez exécuter la tâche SparkR et vous référer à l'image Docker à partir d'Amazon ECR.

Après avoir lancé le cluster, utilisez SSH pour vous connecter à un nœud principal et exécutez les commandes suivantes pour générer l'image Docker locale à partir de l'exemple de Dockerfile SparkR.

Tout d'abord, créez un répertoire et un fichier Dockerfile.

```
mkdir sparkr
vi sparkr/Dockerfile
```

Collez le contenu du fichier Dockerfile SparkR et exécutez les commandes suivantes pour créer une image Docker.

```
sudo docker build -t local/sparkr-example sparkr/
```

Marquez et téléchargez l'image créée localement sur Amazon ECR, en la remplaçant par votre point de *123456789123.dkr.ecr.us-east-1.amazonaws.com* terminaison Amazon ECR.

```
sudo docker tag local/sparkr-example 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-
docker-examples:sparkr-example
sudo docker push 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:sparkr-example
```

Utilisez SSH pour vous connecter au nœud primaire et préparer un script R avec le nom `sparkR.R`. Collez le contenu suivant dans le fichier `sparkR.R`.

```
library(SparkR)
sparkR.session(appName = "R with Spark example", sparkConfig =
 list(spark.some.config.option = "some-value"))

sqlContext <- sparkRSQL.init(spark.sparkContext)
library(randomForest)
# check release notes of randomForest
rfNews()

sparkR.session.stop()
```

Sur Amazon EMR 6.0.0, indiquez le nom de l'image Docker pour soumettre la tâche. Définissez les paramètres de configuration supplémentaires pour vous assurer que l'exécution de la tâche utilise Docker comme environnement d'exécution. Lors de l'utilisation d'Amazon ECR, `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG` doit référencer le fichier `config.json` contenant les informations d'identification utilisées pour s'authentifier auprès d'ECR.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:sparkr-example
DOCKER_CLIENT_CONFIG=hdfs:///user/hadoop/config.json
```

```
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
sparkR.R
```

Sur Amazon EMR 6.1.0 et versions ultérieures, indiquez le nom de l'image Docker pour soumettre la tâche. Lorsque l'authentification automatique ECR est activée, exécutez la commande suivante.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:sparkr-example
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
sparkR.R
```

Lorsque la tâche est terminée, prenez note de l'ID de l'application YARN et utilisez la commande suivante pour obtenir la sortie de la tâche SparkR. Cet exemple inclut des tests permettant de vérifier que la bibliothèque randomForest, la version installée et les notes de mise à jour sont disponibles.

```
yarn logs --applicationId application_id | grep -B4 -A10 "Type rfNews"
randomForest 4.6-14
Type rfNews() to see new features/changes/bug fixes.
Wishlist (formerly TODO):

* Implement the new scheme of handling classwt in classification.

* Use more compact storage of proximity matrix.

* Allow case weights by using the weights in sampling?

=============================================================================
Changes in 4.6-14:
```

Utiliser le catalogue AWS Glue Data Catalog avec Spark sur Amazon EMR

À l'aide d'Amazon EMR version 5.8.0 ou ultérieure, vous pouvez configurer Spark pour utiliser le catalogue de données AWS Glue comme métastore Apache Hive. Nous recommandons cette configuration lorsque vous avez besoin d'un métastore Hive persistant ou d'un métastore Hive partagé par différents clusters, services, applications ou comptes. AWS

À l'aide d'Amazon EMR version 6.5.0 ou ultérieure, vous pouvez configurer Spark pour utiliser le catalogue de données AWS Glue avec Apache Iceberg.

À l'aide d'Amazon EMR version 7.5.0 ou ultérieure, vous pouvez configurer Spark pour utiliser le catalogue de données AWS Glue comme catalogue Iceberg REST.

AWS Glue est un service d'extraction, de transformation et de chargement (ETL) entièrement géré qui permet de classer vos données de manière simple et rentable, de les nettoyer, de les enrichir et de les déplacer de manière fiable entre différents magasins de données. Le catalogue de données AWS Glue fournit un référentiel de métadonnées unifié couvrant une variété de sources de données et de formats de données, s'intégrant à Amazon EMR ainsi qu'à Amazon RDS, Amazon Redshift, Redshift Spectrum, Athena et à toute application compatible avec le métastore Apache Hive. AWS Les robots d'exploration Glue peuvent automatiquement déduire un schéma à partir des données sources dans Amazon S3 et stocker les métadonnées associées dans le catalogue de données. Pour plus d'informations sur le catalogue de données, consultez la section Remplissage du catalogue de données AWS Glue dans le AWS Glue Developer Guide.

Des frais distincts s'appliquent pour AWS Glue. Il existe un tarif mensuel pour le stockage et l'accès aux métadonnées du catalogue de données, un taux horaire facturé par minute pour les tâches AWS Glue ETL et le temps d'exécution des robots d'exploration, et un taux horaire facturé par minute pour chaque point de terminaison de développement provisionné. Le catalogue de données vous permet de stocker jusqu'à un million d'objets sans aucuns frais. Au-delà d'un million d'objets, des frais de 1 USD vous seront facturés tous les 100 000 objets. Dans le catalogue de données, sont considérés comme des objets une table, une partition ou une base de données. Pour de plus amples informations, consultez Tarification Glue.

> **Important**
>
> Si vous avez créé des tables à l'aide d'Amazon Athena ou d'Amazon Redshift Spectrum avant le 14 août 2017, les bases de données et les tables sont stockées dans un catalogue

géré par Athena, distinct du catalogue de données Glue. AWS Pour intégrer Amazon EMR à ces tables, vous devez passer au catalogue de données AWS Glue. Pour plus d'informations, consultez la section Mise à niveau vers le catalogue de données AWS Glue dans le guide de l'utilisateur Amazon Athena.

Spécifier le catalogue de données AWS Glue comme métastore Apache Hive

Vous pouvez spécifier le catalogue de données AWS Glue comme métastore à l'aide de l'API AWS Management Console AWS CLI, ou Amazon EMR. Lorsque vous utilisez la CLI ou l'API, vous utilisez la classification de configuration de Spark pour spécifier le catalogue de données. En outre, avec Amazon EMR 5.16.0 et versions ultérieures, vous pouvez utiliser la classification de configuration pour spécifier un catalogue de données dans un autre. Compte AWS Lorsque vous utilisez la console, vous pouvez spécifier le catalogue de données à l'aide des Options avancées ou des Options rapides.

> **Note**
>
> La possibilité d'utiliser AWS Glue Data Catalog est également disponible avec Zeppelin car Zeppelin est installé avec des composants Spark.

Console

Pour spécifier AWS Glue Data Catalog comme métastore Apache Hive avec la nouvelle console

1. Connectez-vous au et ouvrez la AWS Management Console console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com
2. Sous Amazon EMR activé EC2 dans le volet de navigation de gauche, choisissez Clusters, puis Create cluster.
3. Sous Solution groupée d'application, choisissez Spark ou Custom. Si vous personnalisez votre cluster, assurez-vous de sélectionner Zeppelin ou Spark comme l'une de vos applications.
4. Sous Paramètres du catalogue de données AWS Glue, cochez la case Utiliser pour les métadonnées de table Spark.
5. Choisissez toutes les autres options qui s'appliquent à votre cluster.

6. Pour lancer cluster, choisissez Créer un cluster.

AWS CLI

Pour spécifier le catalogue de données AWS Glue en tant que métastore Apache Hive avec le AWS CLI

Pour plus d'informations sur la spécification d'une classification de configuration à l'aide de l'API AWS CLI et Amazon EMR, consultez. Configuration des applications

- Spécifiez la valeur de `hive.metastore.client.factory.class` à l'aide de la classification `spark-hive-site` comme illustré dans l'exemple suivant :

```
[
  {
    "Classification": "spark-hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory"
    }
  }
]
```

Pour spécifier un catalogue de données dans un autre AWS compte, ajoutez la `hive.metastore.glue.catalogid` propriété comme indiqué dans l'exemple suivant. Remplacez *`acct-id`* par le compte AWS du catalogue de données.

```
[
  {
    "Classification": "spark-hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

Spécifier le catalogue de données AWS Glue en tant que catalogue Apache Iceberg

Vous pouvez spécifier le catalogue de données AWS Glue comme implémentation du catalogue Apache Iceberg, ou comme point de terminaison du catalogue REST Apache Iceberg, à l'aide de l'API AWS Management Console AWS CLI, ou de l'API Amazon EMR, ou lors de la configuration d'exécution de session Spark. Lorsque vous utilisez la CLI ou l'API, vous utilisez la classification de configuration de Spark pour spécifier le catalogue de données. Pour plus de détails, voir Spécifier le catalogue de données AWS Glue comme catalogue Apache Iceberg.

Autorisations IAM

Le profil d' EC2 instance d'un cluster doit disposer des autorisations IAM pour les actions AWS Glue. En outre, si vous activez le chiffrement pour les objets AWS Glue Data Catalog, le rôle doit également être autorisé à chiffrer, déchiffrer et générer les objets AWS KMS key utilisés pour le chiffrement.

Autorisations pour les actions AWS Glue

Si vous utilisez le profil d' EC2 instance par défaut pour Amazon EMR, aucune action n'est requise. La politique `AmazonElasticMapReduceforEC2Role` gérée attachée au `EMR_EC2_DefaultRole` autorise toutes les actions AWS Glue nécessaires. Toutefois, si vous spécifiez un profil d' EC2 instance personnalisé et des autorisations, vous devez configurer les actions AWS Glue appropriées. Utilisez la stratégie gérée `AmazonElasticMapReduceforEC2Role` comme point de départ. Pour plus d'informations, consultez la section Rôle de service pour les EC2 instances de cluster (profil d'EC2 instance) dans le guide de gestion Amazon EMR.

Autorisations pour chiffrer et déchiffrer le catalogue de données Glue AWS

Votre profil d'instance a besoin d'autorisations pour chiffrer et déchiffrer les données au moyen de votre clé. Vous n'avez pas besoin de configurer ces autorisations si les deux déclarations suivantes s'appliquent :

- Vous activez le chiffrement des objets du catalogue de données AWS Glue à l'aide de clés gérées pour AWS Glue.
- Vous utilisez un cluster Compte AWS identique au AWS Glue Data Catalog.

Dans le cas contraire, vous devez ajouter l'instruction suivante à la politique d'autorisation attachée à votre profil d' EC2 instance.

```
[
    {
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "kms:Decrypt",
                    "kms:Encrypt",
                    "kms:GenerateDataKey"
                ],
                "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
            }
        ]
    }
]
```

Pour plus d'informations sur le chiffrement du catalogue de données AWS Glue, consultez la section Chiffrer votre catalogue de données dans le guide du développeur de AWS Glue.

Autorisations basées sur les ressources

Si vous utilisez AWS Glue conjointement avec Hive, Spark ou Presto dans Amazon EMR, AWS Glue prend en charge les politiques basées sur les ressources afin de contrôler l'accès aux ressources du catalogue de données. Ces ressources comprennent les bases de données, les tables, les connexions et les fonctions définies par l'utilisateur. Pour plus d'informations, consultez Politiques de ressources AWS Glue dans le Guide du développeur AWS Glue.

Lorsque vous utilisez des politiques basées sur les ressources pour limiter l'accès à AWS Glue depuis Amazon EMR, le principal que vous spécifiez dans la politique d'autorisation doit être l'ARN du rôle associé au profil d' EC2 instance spécifié lors de la création d'un cluster. Par exemple, pour une politique basée sur les ressources attachée à un catalogue, vous pouvez spécifier l'ARN du rôle de service par défaut pour les EC2 instances de cluster, *EMR_EC2_DefaultRole* en utilisant le format illustré dans l'exemple suivant : `Principal`

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

Il *acct-id* peut être différent de l'identifiant du compte AWS Glue. Cela permet d'accéder aux clusters EMR à partir de comptes différents. Vous pouvez spécifier plusieurs principaux, chacun provenant d'un compte différent.

Considérations relatives à l'utilisation du catalogue de données AWS Glue

Tenez compte des éléments suivants lorsque vous utilisez AWS Glue Data Catalog comme métastore Apache Hive avec Spark :

- Le manque de base de données par défaut sans URI d'emplacement entraîne des échecs lorsque vous créez une table. Pour contourner ce problème, utilisez la clause `LOCATION` pour spécifier un emplacement de bucket, tel que `s3://amzn-s3-demo-bucket1`, lorsque vous utilisez `CREATE TABLE`. Vous pouvez aussi créer des tables dans une autre base de données que celle par défaut.
- Il n'est pas possible de renommer des tables depuis AWS Glue.
- Lorsque vous créez une table Hive sans spécifier de `LOCATION`, les données de la table sont stockées à l'emplacement spécifié par la propriété `hive.metastore.warehouse.dir`. Par défaut, il s'agit d'un emplacement dans HDFS. Si un autre cluster doit accéder à la table, il échoue, sauf s'il dispose d'autorisations adéquates sur le cluster qui a créé la table. En outre, le stockage dans HDFS étant transitoire, si le cluster est résilié, les données de la table sont perdues et la table doit être recréée. Nous vous recommandons de spécifier un `LOCATION` dans Amazon S3 lorsque vous créez une table Hive à l'aide de AWS Glue. Vous pouvez également utiliser la classification de configuration `hive-site` pour spécifier un emplacement dans Amazon S3 pour `hive.metastore.warehouse.dir`, qui s'applique à toutes les tables Hive. Si une table est créée dans un emplacement HDFS et que le cluster qui l'a créée est toujours en cours d'exécution, vous pouvez mettre à jour l'emplacement de la table vers Amazon S3 depuis AWS Glue. Pour plus d'informations, consultez Working with Tables on the AWS Glue Console dans le AWS Glue Developer Guide.
- Les valeurs de partition contenant des guillemets et des apostrophes ne sont pas prises en charge, par exemple `PARTITION (owner="Doe's").`.
- Les statistiques sur les colonnes sont prises en charge à partir de la version emr-5.31.0.
- L'utilisation de l'autorisation Hive n'est pas prise en charge. Comme alternative, envisagez d'utiliser les politiques basées sur les ressources AWS Glue. Pour plus d'informations, consultez Utiliser des politiques basées sur les ressources pour Amazon EMR Access to Glue AWS Data Catalog.

Tenez compte des points suivants lorsque vous utilisez AWS Glue Data Catalog en tant que catalogue REST Apache Iceberg avec Spark :

- Si vous utilisez le catalogue de sessions Spark avec Iceberg, décrit dansDifférences de configuration entre Iceberg SparkCatalog et SparkSessionCatalog, vous devez configurer le catalogue de données AWS Glue en tant que métastore Apache Hive, en plus de configurer le catalogue de données AWS Glue en tant que catalogue REST Apache Iceberg.
- Le point de terminaison IRC AWS Glue Data Catalog prend uniquement en charge le schéma d'authentification Amazon SigV4. OAuth n'est pas pris en charge. Pour OAuth les utilisateurs, veuillez utiliser IAM Identity Center pour configurer l'accès. Voir Connecting Lake Formation with IAM Identity Center.
- Le catalogue REST de AWS Glue Iceberg ne prend pas en charge toutes les opérations en open source.

Utilisation d'une hiérarchie de plusieurs catalogues dans AWS Glue Data Catalog avec Spark sur Amazon EMR

Vous pouvez enregistrer votre cluster Amazon EMR pour accéder au catalogue de données AWS Glue, qui met les tables et autres ressources du catalogue à la disposition de différents consommateurs. AWS Glue Data Catalog prend en charge une hiérarchie de catalogues multiples, qui unifie vos données dans les lacs de données Amazon S3. Il fournit également une API de métastore Hive et une API REST Apache Iceberg open source pour accéder aux données. Ces fonctionnalités sont disponibles pour Amazon EMR et d'autres services tels Amazon Athena qu'Amazon Redshift.

Comment les ressources du catalogue sont organisées

Lorsque vous créez des ressources dans le catalogue de données AWS Glue, vous pouvez y accéder depuis n'importe quel moteur SQL compatible avec l'API REST d'Apache Iceberg ou le métastore Hive. AWS Lake Formation gère les autorisations.

Dans AWS Glue Data Catalog, les données sont organisées selon une hiérarchie logique de catalogues, de bases de données et de tables :

- Catalogue : conteneur logique qui contient des objets provenant d'un magasin de données, tels que des schémas ou des tables.
- Catalogue pour stocker les tables Redshift Managed Storage (RMS) : lorsque vous gérez des catalogues pour stocker des tables RMS, vous pouvez accéder à ces tables à l'aide d'Iceberg.

- Base de données : organise les objets de données tels que les tables et les vues dans un catalogue.
- Tables et vues : objets de données d'une base de données qui fournissent une couche d'abstraction avec un schéma compréhensible. Ils fournissent une couche permettant d'accéder aux données sous-jacentes, qui peuvent être dans différents formats et à différents emplacements.

Configuration d'un catalogue de données à utiliser avec Amazon EMR

Pour commencer, vous configurez le catalogue pour qu'il prenne en charge les outils Amazon EMR. Le catalogue de données AWS Glue est compatible avec le métastore Hive et avec Iceberg REST. APIs

Configuration d'Amazon EMR avec un metastore Hive

Pour plus d'informations sur la configuration de cette option, consultez la section relative à la prise en charge des tâches Spark par AWS Glue Data Catalog dans le guide de l'utilisateur de AWS Glue. Cette rubrique explique comment configurer AWS Glue Data Catalog en tant que métastore Hive et comment le rendre disponible en tant que point de terminaison. De plus, la documentation Amazon EMR vous montre comment spécifier Glue Data Catalog en tant que métastore Spark, dans Use the AWSAWS Glue Data Catalog en tant que métastore Apache Hive pour Spark.

Autorisations d'accès aux ressources dans AWS Glue Data Catalog

Cette section décrit les exigences de la politique IAM relatives à l'utilisation des outils Amazon EMR avec les données du catalogue. Après avoir enregistré votre cluster dans le catalogue de données AWS Glue, vous devez disposer des autorisations suivantes pour découvrir la création et les modifications apportées au catalogue de données créé ultérieurement :

- colle : GetCatalog
- colle : GetCatalogs
- sièges : AssumeRole
- sièges : TagSession
- sièges : SetContext
- sièges : SetSourceIdentity

Dans la plupart des cas, lorsque vous attribuez des autorisations, nous vous recommandons de créer un rôle IAM et de lui attribuer des autorisations.

En outre, pour interroger les données du catalogue, vous devez définir des autorisations pour le catalogue de données à l'aide de AWS Lake Formation. Pour plus d'informations sur la définition des autorisations pour les catalogues de données dans AWS Lake Formation, consultez la section Octroi et révocation d'autorisations sur les ressources du catalogue de données.

Après avoir créé et configuré votre cluster, et défini les autorisations sur les objets de votre catalogue, vous pouvez soumettre des tâches pour interroger et traiter des données.

Configurer Spark pour accéder à une hiérarchie de catalogues multiples dans AWS Glue Data Catalog

Avec EMR 7.5, vous pouvez configurer Spark pour utiliser la hiérarchie multi-catalogues de AWS Glue. Une hiérarchie de catalogues multiples vous permet de :

- Importez vos données de stockage géré Redshift (RMS), telles que les tables, les vues et les vues matérialisées provenant d'entrepôts de données Amazon Redshift existants, vers Glue Data Catalog. AWS Vous pouvez interroger ces objets en utilisant EMR on et EMR EC2 Serverless.
- Créez des catalogues RMS, AWS Glue Data Catalog et stockez les données dans RMS à l'aide de ZeroETL et interrogez les données à l'aide de moteurs de requête compatibles avec Iceberg.
- Créez des tables Iceberg gérées dans AWS Glue Data Catalog grâce à une gestion du stockage complète qui inclut le compactage, les instantanés et la rétention.

Connexion à un catalogue multiple lors de l'initialisation d'une session Spark

Les exemples suivants montrent comment utiliser le shell interactif Spark, Spark Submit ou Amazon EMR Notebooks pour utiliser la hiérarchie multi-catalogues AWS de Glue.

spark-shell

1. Connexion au nœud principal à l'aide de SSH Pour plus d'informations, consultez Connexion au nœud principal à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
2. Entrez la commande suivante pour lancer le shell Spark. Pour utiliser la PySpark coque, remplacez-la `spark-shell` par`pyspark`.

```
spark-shell \
```

```
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket/prefix/
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.catalog.my_catalog.glue.id=Glue RMS catalog ID \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

spark-submit

1. Connexion au nœud principal à l'aide de SSH Pour plus d'informations, consultez Connexion au nœud principal à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
2. Entrez la commande suivante pour lancer la session Spark pour Spark.

```
spark-submit \
--conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
--conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
--conf spark.sql.catalog.my_catalog.type=glue \
--conf spark.sql.catalog.my_catalog.glue.id=Glue RMS catalog ID \
--conf spark.sql.defaultCatalog=my_catalog \
--conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

EMR Studio notebooks

Pour initialiser une session Spark à l'aide des blocs-notes EMR Studio, configurez votre session Spark à l'aide de la commande magique `%%configure` de votre bloc-notes Amazon EMR, comme dans l'exemple suivant. Pour plus d'informations, consultez Utilisation des magies de Blocs-notes EMR dans le Guide de gestion Amazon EMR.

```
%%configure -f{
"conf":{
    "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
    "spark.sql.catalog.my_catalog.type":"glue",
    "spark.sql.catalog.my_catalog.glue.id":"Glue RMS catalog ID",
    "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
    "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
```

```
}
```

CLI

Pour initialiser une session Spark à l'aide de la CLI, exécutez l'exemple suivant. Pour plus d'informations sur la spécification d'une classification de configuration à l'aide de l'API Amazon EMR AWS CLI et Amazon EMR, consultez Configurer les applications.

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
      "spark.sql.catalog.my_catalog.type":"glue",
      "spark.sql.catalog.my_catalog.glue.id":"Glue RMS catalog ID",
      "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
      "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
  }
]
```

Initialisation d'une session Spark dans Redshift Managed Storage with AWS Glue Data Catalog

L'exemple de commande suivant initialise une session Spark avec le AWS Glue Data Catalog.

```
spark-sql \
  --conf spark.sql.catalog.rms=org.apache.iceberg.spark.SparkCatalog \
  --conf spark.sql.catalog.rms.type=glue \
  --conf spark.sql.catalog.rms.glue.id=Glue RMS catalog ID \
  --conf spark.sql.defaultCatalog=rms
  --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

L'exemple suivant initialise une session Spark à l'aide de l'API REST d'Iceberg et du catalogue de données Redshift Managed Storage with Glue AWS .

```
spark-sql \
  --conf spark.sql.catalog.rms=org.apache.iceberg.spark.SparkCatalog \
```

```
  --conf spark.sql.catalog.rms.type=rest \
  --conf spark.sql.catalog.rms.warehouse=glue RMS catalog ID \
  --conf spark.sql.catalog.rms.uri=Glue endpoint URI/iceberg \
  --conf spark.sql.catalog.rms.rest.sigv4-enabled=true \
  --conf spark.sql.catalog.rms.rest.signing-name=glue \
  --conf spark.sql.defaultCatalog=rms \
  --conf
spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Pour plus d'informations sur l'utilisation d'une hiérarchie multi-catalogues AWS Glue avec Spark Iceberg, consultez Utiliser un cluster Iceberg avec Spark.

Considérations et limites relatives à une configuration à catalogues multiples

- L'utilisation d'une hiérarchie de catalogues multiples avec le métastore Apache Hive n'est pas prise en charge.
- L'utilisation d'une hiérarchie de catalogues multiples avec Apache Iceberg ne permet pas de revenir à la métastore Apache Hive lors de l'utilisation. `SparkSessionCatalog`
- L'EMR sur les EC2 clusters dotés du rôle Runtime ne prend pas en charge la hiérarchie multi-catalogues.
- L'EMR sur les EC2 clusters activés ne prend AWS Lake Formation pas en charge la hiérarchie multi-catalogues.

Configuration de Spark

Vous pouvez configurer Spark sur Amazon EMR à l'aide de classifications de configuration. Pour plus d'informations sur les classifications de configuration, consultez Configuration des applications.

Les classifications de configuration pour Spark sur Amazon EMR incluent :

- **`spark`** – Définit la propriété `maximizeResourceAllocation` sur la valeur true ou false. Lorsque la valeur est true, Amazon EMR configure automatiquement les propriétés `spark-defaults` en fonction de la configuration matérielle du cluster. Pour de plus amples informations, veuillez consulter Utiliser `maximizeResourceAllocation`.
- **`spark-defaults`** – Définit les valeurs dans le fichier `spark-defaults.conf`. Pour plus d'informations, consultez Configuration Spark dans la documentation Spark.

- **`spark-env`** – Définit les valeurs dans le fichier `spark-env.sh`. Pour plus d'informations, consultez Variables d'environnement dans la documentation Spark.
- **`spark-hive-site`** – Définit les valeurs dans le `hive-site.xml` pour Spark.
- **`spark-log4j`** : (Amazon EMR versions 6.7.x et antérieures) définit les valeurs dans le fichier `log4j.properties`. Pour plus d'informations, consultez le fichier log4j.properties.template sur Github.
- **`spark-log4j2`** : (Amazon EMR versions 6.8.0 et supérieures) définit les valeurs du fichier `log4j2.properties`. Pour plus d'informations, consultez le fichier log4j2.properties.template sur Github.
- **`spark-metrics`** – Définit les valeurs dans le fichier `metrics.properties`. Pour les paramètres et plus d'informations, consultez le fichier metrics.properties.template sur Github, et Métriques dans la documentation Spark.

> **Note**
>
> Si vous migrez des charges de travail Spark vers Amazon EMR depuis une autre plateforme, nous vous recommandons de tester vos charges de travail avec le Valeurs Spark par défaut définies par Amazon EMR avant d'ajouter des configurations personnalisées. La plupart des clients constatent une amélioration des performances grâce à nos paramètres par défaut.

Rubriques

- Valeurs Spark par défaut définies par Amazon EMR
- Configuration du récupérateur de mémoire Spark sur Amazon EMR 6.1.0
- Utiliser maximizeResourceAllocation
- Configuration du comportement de mise hors service du nœud
- Variable d' ThriftServer environnement Spark
- Modification des paramètres Spark par défaut
- Migration d'Apache Log4j 1.x vers Log4j 2.x

Valeurs Spark par défaut définies par Amazon EMR

La table suivante illustre comment Amazon EMR définit les valeurs par défaut dans `spark-defaults` qui affectent les applications.

Valeurs Spark par défaut définies par Amazon EMR

Paramètre	Description	Valeur par défaut
`spark.executor.memory`	Quantité de mémoire à utiliser par processus d'exécution. Par exemple, `1g`, `2g`.	Ce paramètre est déterminé par les types d'instance de noyau et de tâche dans le cluster.
`spark.executor.cores`	Le nombre de cœurs à utiliser sur chaque exécuteur.	Ce paramètre est déterminé par les types d'instance de noyau et de tâche dans le cluster.
`spark.dynamicAllocation.enabled`	Lorsque cela est vrai, utilisez l'allocation dynamique des ressources pour augmenter ou réduire le nombre d'exécuteurs enregistrés avec une application en fonction de la charge de travail.	`true` (avec Amazon EMR 4.4.0 et versions ultérieures) **Note** Le service Spark réorganisé est automatiquement configuré par Amazon EMR.
`spark.sql.hive.advancedPartitionPredicatePushdown.enabled`	Lorsque c'est vrai, le transfert avancé des prédicats de partition vers le métastore Hive est activé.	`true`
`spark.sql.hive.stringLikePartitionPredicatePushdown.enabled`	Pousse vers le bas les filtres `startsWith` , `contains` et `endsWith` dans le métastore Hive.	`true`

Paramètre	Description	Valeur par défaut
	Note Glue ne prend pas en charge le push down des prédicats pour `startsWith` , `contains`, ou `endsWith`. Si vous utilisez le métastore Glue et que vous rencontrez des erreurs en raison de l'activation des prédicats pour ces fonctions, définissez cette configuration sur `false`.	

Configuration du récupérateur de mémoire Spark sur Amazon EMR 6.1.0

La définition de configurations personnalisées de récupérateur de mémoire avec `spark.driver.extraJavaOptions` et `spark.executor.extraJavaOptions` entraîne un échec du lancement du pilote ou de l'exécuteur avec Amazon EMR 6.1 en raison d'une configuration de récupérateur de mémoire conflictuelle avec Amazon EMR 6.1.0. Pour Amazon EMR 6.1.0, la configuration de récupérateur de mémoire par défaut est définie via `spark.driver.defaultJavaOptions` et `spark.executor.defaultJavaOptions`. Cette configuration s'applique uniquement à Amazon EMR 6.1.0. Les options de la JVM non liées au récupérateur de mémoire, telles que celles permettant de configurer le journal (`-verbose:class`), peuvent toujours être définies via `extraJavaOptions`. Pour plus d'informations, consultez Propriétés de l'application Spark.

Utiliser `maximizeResourceAllocation`

Pour configurer vos exécuteurs afin qu'ils utilisent le maximum de ressources possibles sur chaque nœud d'un cluster, définissez `maximizeResourceAllocation` sur `true` dans votre classification de configuration `spark`. `maximizeResourceAllocation` est spécifique à Amazon EMR. Lorsque vous activez `maximizeResourceAllocation`, Amazon EMR calcule les ressources de calcul et de mémoire maximales disponibles pour un exécuteur sur une instance du groupe d'instances principales. Il définit ensuite les paramètres correspondants `spark-defaults` en fonction des valeurs maximales calculées.

Amazon EMR calcule les ressources de calcul et de mémoire maximales disponibles pour un exécuteur en fonction d'un type d'instance issu du parc d'instances principal. Étant donné que chaque parc d'instances peut avoir des types et des tailles d'instances différents au sein d'un même parc, la configuration d'exécuteur utilisée par Amazon EMR n'est peut-être pas la meilleure pour vos clusters. Nous vous déconseillons donc d'utiliser les paramètres par défaut lors de l'utilisation d'une allocation de ressources maximale. Configurez des paramètres personnalisés pour vos clusters de flotte d'instances.

> **Note**
>
> Vous ne devez pas utiliser l'`maximizeResourceAllocation`option sur des clusters contenant d'autres applications distribuées telles que HBase. Amazon EMR utilise des configurations YARN personnalisées pour les applications distribuées, qui peuvent entrer en conflit avec `maximizeResourceAllocation` et faire échouer les applications Spark.

Voici un exemple de classification de configuration Spark avec `maximizeResourceAllocation` défini sur `true`.

```
[
  {
    "Classification": "spark",
    "Properties": {
      "maximizeResourceAllocation": "true"
    }
  }
]
```

Paramètres configurés dans **`spark-defaults`** lorsque **`maximizeResourceAllocation`** est activé

Paramètre	Description	Valeur
spark.default.parallelism	Le nombre de partitions par défaut RDDs est renvoyé par des transformations telles que jointure et paralléli sation lorsqu'elles ne sont pas définies par l'utilisateur. reduceByKey	2X nombre de cœurs de processeurs disponibles sur des conteneurs YARN.
spark.driver.memory	Quantité de mémoire à utiliser pour le processus du pilote, c'est-à-dire où SparkContext est initialisé. (par exemple, 1 g, 2 g).	Le paramétrage est configuré en fonction des types d'instanc es dans le cluster. Cependant , parce que l'application de pilote Spark peut s'exécute r soit sur le maître, soit sur une des instances principales (par exemple, dans le client YARN et les modes cluster, respectivement), la définitio n dépend du plus petit des types d'instance dans ces deux groupes d'instances.
spark.executor.memory	Volume de mémoire à utiliser par les processus de l'exécute ur. (par exemple, 1g, 2g)	Le paramétrage est configuré en fonction des types d'instanc es de noyau et de tâches dans le cluster.
spark.executor.cores	Le nombre de cœurs à utiliser sur chaque exécuteur.	Le paramétrage est configuré en fonction des types d'instanc es de noyau et de tâches dans le cluster.

Paramètre	Description	Valeur
spark.executor.instances	Le nombre d'exécuteurs.	Le paramétrage est configuré en fonction des types d'instances de noyau et de tâches dans le cluster. Définissez, sauf si `spark.dynamicAllocation.enabled` définit explicitement sur true en même temps.

Configuration du comportement de mise hors service du nœud

Lorsque vous utilisez Amazon EMR 5.9.0 ou une version ultérieure, Spark sur Amazon EMR inclut un ensemble de fonctionnalités qui permettent de s'assurer que Spark gère plus élégamment l'arrêt des nœuds suite à un redimensionnement manuel ou à une demande de politique de mise à l'échelle automatique. Amazon EMR implémente un mécanisme de liste de refus dans Spark par-dessus le mécanisme de mise hors service de YARN. Ce mécanisme aide à garantir qu'aucune nouvelle tâche n'est planifiée sur un nœud qui est mis hors service, tout en autorisant dans le même temps que les tâches déjà en cours d'exécution se terminent. En outre, il existe des fonctionnalités qui permettent de rétablir les tâches Spark plus rapidement en cas de perte de blocs aléatoires lors de la terminaison d'un nœud. Le processus de recalcul est déclenché plus tôt et optimisé pour un recalcul plus rapide avec moins de tentatives d'étape ; en outre, il est possible d'empêcher l'échec des tâches dû aux défaillances d'extraction provoquées par les blocs aléatoires manquants.

> **Important**
>
> Le paramètre `spark.decommissioning.timeout.threshold` a été ajouté dans la version 5.11.0 d'Amazon EMR pour améliorer la résilience de Spark lorsque vous utilisez des instances Spot. Dans les versions précédentes, lorsqu'un nœud utilise une instance Spot et que l'instance est terminée en raison du prix de l'offre, Spark peut ne pas gérer correctement la terminaison. Les tâches peuvent échouer et les recalculs aléatoires peuvent nécessiter un temps important. Pour cette raison, nous vous conseillons d'utiliser la version 5.11.0 ou une version ultérieure si vous utilisez des instances Spot.

Paramètres de mise hors service d'un nœud Spark

Paramètre	Description	Valeur par défaut
`spark.blacklist.decommissioning.enabled`	Lorsqu'il est défini sur `true`, Spark refuse de répertorier les nœuds qui sont dans l'état `decommissioning` dans YARN. Spark ne planifie pas de nouvelles tâches sur les exécuteurs s'exécutant sur ce nœud. Les tâches déjà en cours d'exécution sont autorisées à se terminer.	`true`
`spark.blacklist.decommissioning.timeout`	La durée pendant laquelle un nœud dans l'état `decommissioning` est sur la liste de refus. Par défaut, cette valeur est définie sur une heure, qui est aussi l'heure par défaut pour `yarn.resourcemanager.decommissioning.timeout`. Pour s'assurer qu'un nœud est placé sur liste noire pendant toute la période de sa mise hors service, définissez une valeur égale ou supérieure à `yarn.resourcemanager.decommissioning.timeout`. Une fois le délai de mise hors service expiré, le nœud passe à un `decommissioned` état et Amazon EMR peut mettre fin à l'instance du nœud. EC2 Si des tâches continuent à	`1h`

Paramètre	Description	Valeur par défaut
	s'exécuter après l'expiration du délai, elles sont perdues ou supprimées, puis replanifiées sur les exécuteurs s'exécutant sur d'autres nœuds.	
`spark.decommissioning.timeout.threshold`	Disponible dans Amazon EMR version 5.11.0 ou ultérieure. Indiqué en secondes. Lorsqu'un nœud passe à l'état de mise hors service, si l'hôte est mis hors service au cours d'une durée inférieure ou égale à cette valeur, Amazon EMR placera le nœud sur la liste noire et nettoiera l'état de l'hôte (comme indiqué par `spark.resourceManager.cleanupExpiredHost`) sans attendre que le nœud passe à l'état de mise hors service. Cela permet à Spark de mieux traiter les arrêts d'instances Spot car les instances Spot sont mises hors services dans un délai de 20 secondes, et ce quelle que soit la valeur de `yarn.resourcemager.decommissioning.timeout`, qui peut ne pas fournir aux autres nœuds suffisamment de temps pour lire les fichiers de lecture aléatoire.	`20s`

Paramètre	Description	Valeur par défaut
`spark.resourceManager.cleanupExpiredHost`	Lorsque la valeur est `true`, Spark annule l'inscription de l'ensemble des données mises en cache et des blocs aléatoires stockés dans les exécuteurs des nœuds qui se trouvent à l'état `decommissioned` . Le processus de récupération s'en trouve accéléré.	`true`
`spark.stage.attempt.ignoreOnDecommissionFetchFailure`	Lorsque la valeur est `true`, aide à empêcher que les phases Spark n'échouent et n'entraînent l'échec de la tâche en raison du trop grand nombre d'extractions à partir des nœuds mis hors service ayant elles-mêmes échoué. Les extractions défaillantes de blocs aléatoires d'un nœud ayant l'état `decommissioned` ne sont pas comptabilisées dans le nombre maximal d'extractions défaillantes consécutives.	`true`

Variable d' ThriftServer environnement Spark

Spark définit la variable d'environnement du port de serveur Thrift Hive, `HIVE_SERVER2_THRIFT_PORT`, sur 10001.

Modification des paramètres Spark par défaut

Vous modifiez les paramètres par défaut dans `spark-defaults.conf` à l'aide de la classification de configuration `spark-defaults` ou du paramètre `maximizeResourceAllocation` dans la classification de configuration `spark`.

Les procédures suivantes montrent comment modifier les paramètres à l'aide de l'interface de ligne ou de la console.

Pour créer un cluster avec spark.executor.memory défini sur 2g à l'aide de la CLI

- Créez un cluster avec Spark installé et la valeur `spark.executor.memory` définie sur 2g, à l'aide de la commande suivante, qui fait référence à un fichier, `myConfig.json` stocké dans Amazon S3.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Spark \
--instance-type m5.xlarge --instance-count 2 --service-role EMR_DefaultRole_V2
 --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

`myConfig.json`:

```
[
    {
      "Classification": "spark-defaults",
      "Properties": {
        "spark.executor.memory": "2G"
      }
    }
  ]
```

Pour créer un cluster avec spark.executor.memory défini sur 2g à l'aide de la console

1. Accédez à la nouvelle console Amazon EMR et sélectionnez Changer pour l'ancienne console depuis le menu latéral. Pour plus d'informations sur ce qu'implique le passage à l'ancienne console, consultez la rubrique Utilisation de l'ancienne console.
2. Choisissez Créer un cluster et Go to advanced options (Aller aux options avancées).
3. Choisissez Spark.
4. Sous Edit software settings (Modifier les paramètres logiciels), conservez l'option Enter configuration (Saisir la configuration) et saisissez la configuration suivante :

```
classification=spark-defaults,properties=[spark.executor.memory=2G]
```

5. Sélectionnez d'autres options, choisissez , puis Create cluster (Créer un cluster).

À régler maximizeResourceAllocation

- Créez un cluster avec Spark installé et `maximizeResourceAllocation` défini sur true en utilisant le AWS CLI, en faisant référence à un fichier`myConfig.json`, stocké dans Amazon S3.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Spark \
--instance-type m5.xlarge --instance-count 2 --service-role EMR_DefaultRole_V2
 --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

`myConfig.json`:

```
[
  {
    "Classification": "spark",
    "Properties": {
      "maximizeResourceAllocation": "true"
```

```
    }
  }
]
```

> **Note**
>
> Avec la version 5.21.0 et ultérieures d'Amazon EMR, vous permet de remplacer les configurations de cluster et de spécifier des classifications de configuration supplémentaires pour chaque groupe d'instances dans un cluster en cours d'exécution. Pour ce faire, utilisez la console Amazon EMR, le AWS Command Line Interface (AWS CLI) ou le AWS SDK. Pour plus d'informations, consultez Fourniture d'une configuration pour un groupe d'instances dans un cluster en cours d'exécution.

Migration d'Apache Log4j 1.x vers Log4j 2.x

Les versions 3.2.x et antérieures d'Apache Spark utilisent l'ancien Apache Log4j 1.x et le fichier `log4j.properties` pour configurer Log4j dans les processus Spark. Les versions 3.3.0 et ultérieures d'Apache Spark utilisent Apache Log4j 2.x et le fichier `log4j2.properties` pour configurer Log4j dans les processus Spark.

Si vous avez configuré Apache Spark Log4j à l'aide d'une version Amazon EMR inférieure à 6.8.0, vous devez supprimer l'ancienne classification `spark-log4j` de configuration et migrer vers la classification de configuration `spark-log4j2` et le format de clé avant de pouvoir passer à Amazon EMR 6.8.0 ou version ultérieure. L'ancienne classification `spark-log4j` entraîne l'échec de la création de clusters avec une erreur `ValidationException` dans les versions 6.8.0 et ultérieures d'Amazon EMR. Aucun frais ne vous sera facturé en cas de panne liée à l'incompatibilité de Log4j, mais vous devez supprimer la classification de configuration obsolète `spark-log4j` pour continuer.

Pour plus d'informations sur la migration d'Apache Log4j 1.x vers Log4j 2.x, consultez le Guide de migration d'Apache Log4j et le Modèle Spark Log4j 2 sur Github.

> **Note**
>
> Avec Amazon EMR, Apache Spark utilise un fichier `log4j2.properties` plutôt que le fichier .xml décrit dans le Guide de migration d'Apache Log4j. De plus, nous ne recommandons pas d'utiliser la méthode du pont Log4j 1.x pour convertir en Log4j 2.x.

Optimisation des performances de Spark

Amazon EMR propose plusieurs fonctions d'optimisation des performances pour Spark. Cette rubrique explique en détail chaque fonction d'optimisation.

Pour de plus amples informations sur la définition de la configuration Spark, veuillez consulter Configuration de Spark.

Exécution de requêtes adaptative

L'exécution adaptative des requêtes est un framework permettant de réoptimiser les plans de requêtes en fonction des statistiques d'exécution. À partir de la version 5.30.0 d'Amazon EMR, les optimisations d'exécution adaptative des requêtes suivantes d'Apache Spark 3 sont disponibles sur Apache Amazon EMR Runtime pour Spark 2.

- Conversion de jointure adaptative
- Coalescence adaptative de partitions shuffle

Conversion de jointure adaptative

La conversion adaptative des jointures améliore les performances des requêtes en convertissant les sort-merge-join broadcast-hash-joins opérations en opérations basées sur la taille d'exécution des étapes de requête. Broadcast-hash-joinsont tendance à être plus performantes lorsqu'un côté de la jointure est suffisamment petit pour diffuser efficacement sa sortie sur tous les exécuteurs, évitant ainsi d'avoir à mélanger les échanges et à trier les deux côtés de la jointure. La conversion adaptative des jointures élargit l'éventail des cas dans lesquels Spark s'exécute broadcast-hash-joins automatiquement.

Cette caractéristique est activée par défaut. Il peut être désactivé en le définissant `spark.sql.adaptive.enabled` sur `false`, ce qui désactive également le cadre d'exécution adaptative des requêtes. Spark décide de sort-merge-join convertir a en a broadcast-hash-join lorsque la statistique de taille d'exécution de l'un des côtés de la jointure ne dépasse pas`spark.sql.autoBroadcastJoinThreshold`, la valeur par défaut étant de 10 485 760 octets (10 MiB).

Coalescence adaptative de partitions réorganisées

La fusion adaptative des partitions de distribution améliore les performances des requêtes en fusionnant de petites partitions de distribution contiguës afin d'éviter la surcharge liée à un trop grand

nombre de petites tâches. Cela vous permet de configurer un plus grand nombre de partitions de shuffle initiales à l'avance, puis de les réduire à une taille ciblée lors de l'exécution, ce qui augmente les chances d'avoir des partitions shuffle réparties de manière plus uniforme.

Cette fonctionnalité est activée par défaut, sauf si `spark.sql.shuffle.partitions` est explicitement défini. Il peut être activé en réglant `spark.sql.adaptive.coalescePartitions.enabled` sur `true`. Le nombre initial de partitions réorganisées et la taille de la partition cible peuvent être ajustés à l'aide des propriétés `spark.sql.adaptive.coalescePartitions.minPartitionNum` et `spark.sql.adaptive.advisoryPartitionSizeInBytes` respectivement. Consultez le tableau suivant pour plus de détails sur les propriétés Spark associées à cette fonctionnalité.

Propriétés de la partition de coalescence adaptative Spark

Propriété	Valeur par défaut	Description
`spark.sql.adaptive.coalescePartitions.enabled`	true, sauf si `spark.sql.shuffle.partitions` est explicitement défini.	Lorsque true et spark.sql.adaptive.enabled est true, Spark fusionne les partitions réorganisées contiguës en fonction de la taille cible (spécifiée par `spark.sql.adaptive.advisoryPartitionSizeInBytes`), afin d'éviter un trop grand nombre de petites tâches.
`spark.sql.adaptive.advisoryPartitionSizeInBytes`	64 Mo	Taille conseillée en octets de partition shuffle lors de la fusion. Cette configuration n'a d'effet que lorsque `spark.sql.adaptive.enabled` et `spark.sql.adaptive.coalescePartitions.enabled` sont tous deux `true`.

Propriété	Valeur par défaut	Description
`spark.sql.adaptive.coalescePartitions.minPartitionNum`	25	Le nombre minimum de partitions shuffle après fusion. Cette configuration n'a d'effet que lorsque `spark.sql.adaptive.enabled` et `spark.sql.adaptive.coalescePartitions.enabled` sont tous deux `true`.
`spark.sql.adaptive.coalescePartitions.initialPartitionNum`	1 000	Nombre initial de partitions réorganisées avant la fusion. Cette configuration n'a d'effet que lorsque `spark.sql.adaptive.enabled` et `spark.sql.adaptive.coalescePartitions.enabled` sont tous deux `true`.

Nettoyage dynamique de partition

Le nettoyage dynamique de partition améliore les performances de tâche avec plus de précision en sélectionnant des partitions spécifiques au sein d'une table qui doivent être lues et traitées pour une requête spécifique. En réduisant la quantité de données lues et traitées, l'exécution du travail est nettement moins longue. Avec Amazon EMR 5.26.0, cette fonctionnalité est activée par défaut. Avec Amazon EMR 5.24.0 et 5.25.0, vous pouvez activer cette fonctionnalité en définissant la propriété Spark `spark.sql.dynamicPartitionPruning.enabled` depuis Spark ou lors de la création de clusters.

Propriétés des partitions d'élimination dynamique Spark

Propriété	Valeur par défaut	Description
`spark.sql.dynamicPartitionPruning.enabled`	`true`	Lorsque c'est vrai, activez l'élimination dynamique des partitions.
`spark.sql.optimizer.dynamicPartitionPruning.enforceBroadcastReuse`	`true`	Si la valeur est `true`, Spark effectue une vérification défensive avant l'exécution de la requête pour s'assurer que la réutilisation des échanges de diffusion dans les filtres d'élimination dynamique n'est pas interrompue par des règles de préparation ultérieures, telles que les règles en colonnes définies par l'utilisateur. Lorsque la réutilisation est interrompue et que cette configuration est `true`, Spark supprime les filtres d'élimination dynamique concernés afin de se prémunir contre les problèmes de performance et d'exactitude. Des problèmes d'exactitude peuvent survenir lorsque l'échange de diffusion du filtre d'élimination dynamique produit des résultats différents et incohérents par rapport à l'échange de diffusion de l'opération de jointure correspondante. La définition de cette configuration sur

Propriété	Valeur par défaut	Description
		`false` doit être effectuée avec prudence ; elle permet de contourner les scénarios, par exemple lorsque la réutilisation est interrompue par des règles en colonnes définies par l'utilisateur. Lorsque l'exécuti on adaptative des requêtes est activée, la réutilisation des diffusions est toujours appliquée.

Cette optimisation améliore les fonctionnalités existantes de Spark 2.4.2, qui prend uniquement en charge les transmission des prédicats statiques qui peuvent être résolus au moment voulu.

Voici des exemples de prédicat statiques en mode push Spark 2.4.2.

```
partition_col = 5

partition_col IN (1,3,5)

partition_col between 1 and 3

partition_col = 1 + 3
```

Le nettoyage dynamique de partition permet au moteur Spark de déduire dynamiquement au moment de l'exécution les partitions qui doivent être lues et qui peuvent être éliminées. Par exemple, la requête suivante implique deux tables : `store_sales` (qui contient toutes les ventes totales pour tous les magasins et est partitionnée par région) et `store_regions` (qui contient un mappage de régions pour chaque pays). Les tables contiennent des données sur les magasins répartis dans le monde entier, mais nous interrogeons uniquement des données pour l'Amérique du Nord.

```
select ss.quarter, ss.region, ss.store, ss.total_sales
from store_sales ss, store_regions sr
where ss.region = sr.region and sr.country = 'North America'
```

Sans nettoyage dynamique de partition, cette requête lit toutes les régions avant de filtrer le sous-ensemble de régions qui correspondent aux résultats de la sous-requête. Avec le nettoyage dynamique de partition, cette requête lit et traite uniquement les partitions pour les régions renvoyées dans la sous-requête. Cela permet d'économiser du temps et des ressources en lisant moins de données dans le stockage et en traitant moins d'enregistrements.

Aplatissement des sous-requêtes scalaires

Cette optimisation améliore les performances des requêtes qui ont des sous-requêtes scalaires sur la même table. Avec Amazon EMR 5.26.0, cette fonctionnalité est activée par défaut. Avec Amazon EMR 5.24.0 et 5.25.0, vous pouvez l'activer en définissant la propriété Spark `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled` depuis Spark ou lors de la création de clusters. Lorsque cette propriété a la valeur true, l'optimiseur de requête aplatit les sous-requêtes scalaires regroupées qui utilisent la même relation si possible. Les sous-requêtes scalaires sont mises à plat en poussant tous les prédicats présents dans la sous-requête dans les fonctions d'agrégation, puis en effectuant une agrégation, avec toutes les fonctions d'agrégation, par relation.

Voici un exemple de requête qui bénéficie de cette optimisation.

```
select (select avg(age) from students                    /* Subquery 1 */
                where age between 5 and 10) as group1,
       (select avg(age) from students                    /* Subquery 2 */
                where age between 10 and 15) as group2,
       (select avg(age) from students                    /* Subquery 3 */
                where age between 15 and 20) as group3
```

L'optimisation réécrit la requête précédente comme suit :

```
select c1 as group1, c2 as group2, c3 as group3
from (select avg (if(age between 5 and 10, age, null)) as c1,
             avg (if(age between 10 and 15, age, null)) as c2,
             avg (if(age between 15 and 20, age, null)) as c3 from students);
```

Notez que la requête réécrite lit la table des étudiants une seule fois et les prédicats des trois sous-requêtes sont ajoutés à la fonction `avg`.

DISTINCT avant INTERSECT

Cette optimisation optimise les jointures lorsque vous utilisez INTERSECT. Avec Amazon EMR 5.26.0, cette fonctionnalité est activée par défaut. Avec Amazon EMR 5.24.0 et 5.25.0, vous pouvez l'activer en définissant la propriété Spark `spark.sql.optimizer.distinctBeforeIntersect.enabled` depuis Spark ou lors de la création de clusters. Les requêtes utilisant INTERSECT sont automatiquement converties pour utiliser une semi-jointure gauche. Lorsque cette propriété est définie sur true, l'optimiseur de requêtes envoie l'opérateur DISTINCT aux enfants d'INTERSECT s'il détecte que l'opérateur DISTINCT peut faire en sorte que le semi-gauche joigne a au lieu de a. BroadcastHashJoin SortMergeJoin

Voici un exemple de requête qui bénéficie de cette optimisation.

```
(select item.brand brand from store_sales, item
     where store_sales.item_id = item.item_id)
intersect
(select item.brand cs_brand from catalog_sales, item
     where catalog_sales.item_id = item.item_id)
```

Sans activer cette propriété `spark.sql.optimizer.distinctBeforeIntersect.enabled`, la requête sera réécrite comme suit.

```
select distinct brand from
  (select item.brand brand from store_sales, item
     where store_sales.item_id = item.item_id)
left semi join
   (select item.brand cs_brand from catalog_sales, item
     where catalog_sales.item_id = item.item_id)
 on brand <=> cs_brand
```

Lorsque vous activez cette propriété `spark.sql.optimizer.distinctBeforeIntersect.enabled`, la requête est réécrite comme suit.

```
select brand from
  (select distinct item.brand brand from store_sales, item
     where store_sales.item_id = item.item_id)
left semi join
   (select distinct item.brand cs_brand from catalog_sales, item
     where catalog_sales.item_id = item.item_id)
```

```
on brand <=> cs_brand
```

Jonction de filtre Bloom

Cette optimisation peut améliorer les performances de certaines jonctions en préfiltrant un côté de la jonction à l'aide d'un filtre Bloom généré à partir des valeurs de l'autre côté de la jonction. Avec Amazon EMR 5.26.0, cette fonctionnalité est activée par défaut. Avec Amazon EMR 5.25.0, vous pouvez activer cette fonctionnalité en définissant la propriété Spark `spark.sql.bloomFilterJoin.enabled` sur `true` depuis Spark ou lors de la création de clusters.

L'exemple suivant est une requête qui peut bénéficier d'un filtre Bloom.

```
select count(*)
from sales, item
where sales.item_id = item.id
and item.category in (1, 10, 16)
```

Lorsque cette fonction est activée, le filtre Bloom est construit à partir de tous les ID d'éléments dont la catégorie figure dans l'ensemble des catégories interrogées. Pendant l'analyse de la table des ventes, le filtre Bloom est utilisé pour déterminer les ventes pour les éléments qui ne sont pas dans l'ensemble défini par le filtre Bloom. Ainsi, ces ventes identifiées peuvent être filtrées dès que possible.

Réorganisation optimisée des jonctions

Cette optimisation peut améliorer les performances des requêtes en réorganisant les jonctions impliquant des tables avec des filtres. Avec Amazon EMR 5.26.0, cette fonctionnalité est activée par défaut. Avec Amazon EMR 5.25.0, vous pouvez activer cette fonctionnalité en définissant le paramètre de configuration Spark `spark.sql.optimizer.sizeBasedJoinReorder.enabled` sur true. Par défaut, Spark procède à la jonction des tables de gauche à droite, comme indiqué dans la requête. Cette stratégie peut ignorer des possibilités d'exécution de jonctions plus petites avec des filtres dans un premier temps, afin de bénéficier des jonctions plus coûteuses ultérieurement.

L'exemple de requête ci-dessous indique tous les éléments renvoyés à partir de tous les magasins d'un pays. Sans réorganisation optimisée des jonctions, Spark joint les deux tables volumineuses `store_sales` et `store_returns` en premier, puis les joint avec `store` et finalement avec `item`.

```
select ss.item_value, sr.return_date, s.name, i.desc,
```

```
from store_sales ss, store_returns sr, store s, item i
where ss.id = sr.id and ss.store_id = s.id and ss.item_id = i.id
and s.country = 'USA'
```

Avec la réorganisation optimisée des jonctions, Spark joint d'abord `store_sales` et `store`, car `store` possède un filtre et est plus petite que `store_returns` et `broadcastable`. Ensuite, Spark joint `store_returns`, puis `item`. Si `item` disposait d'un filtre et pouvait être diffusée, elle pourrait bénéficier d'une réorganisation, ce qui entraînerait la jonction de `store_sales` avec `store`, puis avec `item`, et finalement avec `store_returns`.

Mise en cache des fragments de résultats Spark

Amazon EMR 6.6.0 et versions ultérieures incluent la fonctionnalité optionnelle de mise en cache des fragments de résultats de Spark qui met automatiquement en cache les fragments de résultats. Ces fragments de résultats font partie des résultats issus de sous-arborescences de requêtes stockées dans un compartiment Amazon S3 de votre choix. Les fragments de résultats de requête stockés sont réutilisés lors des exécutions de requêtes suivantes, ce qui permet d'accélérer les requêtes.

La mise en cache des fragments de résultats analyse vos requêtes Spark et met en cache les fragments de résultats éligibles dans l'emplacement S3 que vous avez spécifié. Lors des exécutions de requête suivantes, les fragments de résultats de requête utilisables sont automatiquement détectés et extraits depuis S3. La mise en cache des fragments de résultats est différente de la mise en cache des ensembles de résultats, dans laquelle les requêtes suivantes doivent correspondre exactement à la requête d'origine pour renvoyer des résultats depuis le cache. Lorsqu'elle est utilisée pour des requêtes qui ciblent de manière répétée un sous-ensemble statique de vos données, la mise en cache des fragments de résultats accélère considérablement les performances.

Considérez la requête suivante, qui compte les commandes jusqu'en 2022 :

```
select
    l_returnflag,
    l_linestatus,
    count(*) as count_order
from
    lineitem
where
    l_shipdate <= current_date
    and year(l_shipdate) == '2022'
group by
    l_returnflag,
```

```
    l_linestatus
```

Au fil du temps, cette requête doit être exécutée tous les jours pour indiquer le total des ventes de l'année. Sans la mise en cache des fragments de résultats, les résultats pour tous les jours de l'année devront être recalculés chaque jour. La requête deviendra plus lente au fil du temps et sera plus lente à la fin de l'année, lorsque les 365 jours de résultats devront être recalculés.

Lorsque vous activez la mise en cache des fragments de résultats, vous utilisez les résultats de tous les jours précédents de l'année à partir du cache. Chaque jour, la fonctionnalité ne doit recalculer qu'un seul jour de résultats. Une fois que la fonctionnalité a calculé le fragment de résultat, elle met le fragment en cache. Par conséquent, les temps de requête activés par le cache sont rapides et restent constants pour chaque requête suivante.

Activation de la mise en cache des fragments de résultats Spark

Pour activer la mise en cache des fragments de résultats, veuillez suivre les étapes ci-dessous :

1. Créez un compartiment de cache dans Amazon S3 et autorisez read/write l'accès à EMRFS. Pour de plus amples informations, veuillez consulter Autorisation d'accès aux données EMRFS dans Amazon S3.
2. Configurez Amazon EMR Spark pour activer la fonctionnalité.

   ```
   spark.subResultCache.enabled = true
   spark.subResultCache.fs.root.path = s3://&example-s3-bucket;/cache_dir/
   ```
3. Activez la gestion du cycle de vie S3 pour que le compartiment nettoie automatiquement les fichiers de cache.
4. Configurez éventuellement les maxBufferSize propriétés reductionRationThreshold et pour affiner la fonctionnalité.

   ```
   spark.sql.subResultCache.reductionRatioThreshold
   spark.sql.subResultCache.maxBufferSize
   ```

Considérations relatives à l'utilisation de la mise en cache des fragments de résultats

Les économies réalisées lorsque vous utilisez des résultats déjà mis en cache dans Amazon S3 plutôt que de les recalculer augmentent avec le nombre de fois où les mêmes résultats mis en cache

peuvent être utilisés. Les requêtes comportant des analyses de grandes tables suivies de filtres ou d'agrégations de hachage qui réduisent la taille du résultat d'au moins 8 (c'est-à-dire un ratio d'au moins 8:1 dans la taille d'entrée:résultats) bénéficieront le plus de cette fonctionnalité. Plus le rapport de réduction entre l'entrée et les résultats est élevé, plus le rapport coût-bénéfice est important. Les requêtes présentant des taux de réduction inférieurs, mais comportant des étapes de calcul coûteuses entre l'analyse des tables et le filtrage ou les agrégations, en bénéficieront également, à condition que le coût de production des résultats soit supérieur à celui de leur extraction sur Amazon S3. Par défaut, la mise en cache des fragments de résultats ne prend effet que lorsqu'elle détecte qu'un taux de réduction sera d'au moins 8:1.

Lorsque vos requêtes réutilisent à plusieurs reprises les résultats mis en cache, les avantages de cette fonctionnalité sont les plus importants. Les requêtes de fenêtre progressives et incrémentielles en sont de bons exemples. Par exemple, une requête à fenêtre continue de 30 jours qui s'exécute déjà depuis 29 jours ne devrait extraire que 1/30e des données cibles de sa source d'entrée d'origine et utiliserait des fragments de résultats mis en cache pour les 29 jours précédents. Une requête de fenêtre incrémentielle en bénéficierait encore plus, car le début de la fenêtre reste fixe : à chaque invocation de la requête, un faible pourcentage du traitement nécessitera une lecture depuis la source d'entrée.

Voici d'autres considérations à prendre en compte lors de l'utilisation de la mise en cache des fragments de résultats :

- Les requêtes qui ne ciblent pas les mêmes données avec les mêmes fragments de requête auront un faible taux d'accès au cache et ne bénéficieront donc pas de cette fonctionnalité.
- Les requêtes présentant de faibles ratios de réduction qui ne contiennent pas d'étapes de calcul coûteuses produiront des résultats mis en cache à peu près aussi coûteux à lire qu'à traiter initialement.
- La première requête démontrera toujours une régression mineure en raison du coût d'écriture dans le cache.
- La fonctionnalité de mise en cache des fragments de résultats fonctionne exclusivement avec les fichiers Parquet. Les autres formats de fichiers ne sont pas pris en charge.
- Les tampons de la fonctionnalité de mise en cache des fragments de résultats essaieront uniquement de mettre en cache les scans dont la taille de fractionnement des fichiers est supérieure ou égale à 128 Mo. Avec la configuration par défaut de Spark, la mise en cache des fragments de résultats est désactivée si la taille de numérisation (taille totale de tous les fichiers analysés) divisée par le nombre de cœurs d'exécuteurs est inférieure à 128 Mo. Lorsque l'une

des configurations Spark répertoriées ci-dessous est définie, la taille du fichier fractionné sera la suivante :

```
min(maxPartitionBytes, max(openCostInBytes, scan size / minPartitionNum))
```

  - spark.sql. leafNodeDefaultParallélisme (la valeur par défaut est spark.default.parallelism)
  - fichiers spark.sql. minPartitionNum (la valeur par défaut est spark.sql. leafNodeDefaultParallélisme)
  - fichiers spark.sql. openCostInOctets
  - fichiers spark.sql. maxPartitionBytes
- La fonction de mise en cache des fragments de résultats est mise en cache au niveau de la granularité de la partition RDD. Le ratio de réduction décrit précédemment, qui est par défaut de 8:1, est évalué par partition RDD. Les charges de travail présentant des ratios de réduction par RDD supérieurs ou inférieurs à 8:1 peuvent présenter des avantages en termes de performances moindres que les charges de travail présentant des ratios de réduction par RDD constamment inférieurs à 8:1.
- La fonctionnalité de mise en cache des fragments de résultats utilise une mémoire tampon d'écriture de 16 Mo par défaut pour chaque partition RDD mise en cache. Si plus de 16 Mo sont mis en cache par partition RDD, le coût lié à la détermination de l'impossibilité d'une écriture peut entraîner une régression des performances.
- Bien que, par défaut, la mise en cache des fragments de résultats ne tente pas de mettre en cache les résultats des partitions RDD avec un ratio de réduction inférieur à 8:1 et plafonne sa mémoire tampon d'écriture à 16 Mo, ces deux valeurs sont réglables selon les configurations suivantes :

```
spark.sql.subResultCache.reductionRatioThreshold (default: 8.0)
spark.sql.subResultCache.maxBufferSize (default: 16MB, max: 64MB)
```

- Plusieurs clusters utilisant la même version d’Amazon EMR peuvent partager le même emplacement de cache. Pour garantir l'exactitude des résultats, la mise en cache des fragments de résultats n'utilisera pas les résultats du cache écrits par différentes versions d'Amazon EMR.
- La mise en cache des fragments de résultats sera automatiquement désactivée pour les cas d'utilisation de Spark Streaming ou lorsque RecordServer Apache Ranger AWS Lake Formation est utilisé.
- Le cache de fragments obtenu read/writes utilise les compartiments EMRFS/S3A et Amazon S3. Les chiffrements CSE (uniquement avec EMRFS) /SSE S3/ SSE KMS sont pris en charge. À titre de contexte, S3A fournit une implémentation Hadoop permettant à un cluster de lire et d'écrire

des données vers et depuis Amazon S3. Notez que la prise en charge de S3A est disponible avec EMR-7.4.0 et versions ultérieures.

Utilisation de l'accélérateur Nvidia RAPIDS pour Apache Spark

Avec Amazon EMR version 6.2.0 et versions ultérieures, vous pouvez utiliser le plugin RAPIDS Accelerator for Apache Spark de Nvidia pour accélérer Spark à l'aide de types d'instances d'unités de traitement EC2 graphique (GPU). L'accélérateur RAPIDS accélérera par GPU vos pipelines de science des données Apache Spark 3.0 sans modifier le code et accélérera le traitement des données et l'entraînement des modèles, tout en réduisant considérablement les coûts d'infrastructure.

Les sections suivantes vous guident dans la configuration de votre cluster EMR pour utiliser le plug-in Spark-rapids pour Spark.

Choix des types d'instance

Pour utiliser le plugin Nvidia Spark-rapids pour Spark, les groupes d'instances principaux et de tâches doivent utiliser des types d'instances EC2 GPU répondant aux exigences matérielles de Spark-rapids. Pour la liste complète des types d'instances GPU pris en charge par Amazon EMR, voir la rubrique Types d'instances pris en charge du Guide de gestion d'Amazon EMR. Le type d'instance du groupe d'instances principal peut être du type GPU ou non, mais les types d'instances ARM ne sont pas pris en charge.

Configuration des configurations d'applications pour votre cluster

1. Activer Amazon EMR pour installer les plug-ins sur votre nouveau cluster

Pour installer des plug-ins, fournissez la configuration suivante lors de la création de votre cluster :

```
{
 "Classification":"spark",
 "Properties":{
  "enableSparkRapids":"true"
 }
}
```

2. Configurer YARN pour utiliser le GPU

Pour plus d'informations sur l'utilisation du GPU sur YARN, voir la rubrique Utilisation du GPU sur YARN de la documentation Apache Hadoop. Voici des exemples de configurations YARN pour les versions 6.x et 7.x d'Amazon EMR :

Amazon EMR 7.x

Exemple de configuration YARN pour Amazon EMR 7.x

```
{
    "Classification":"yarn-site",
    "Properties":{
        "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
        "yarn.resource-types":"yarn.io/gpu",
        "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
        "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/usr/
bin",
        "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
        "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/spark-
rapids-cgroup",
        "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
        "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
    }
},{
    "Classification":"container-executor",
    "Properties":{

    },
    "Configurations":[
        {
            "Classification":"gpu",
            "Properties":{
                "module.enabled":"true"
            }
        },
        {
            "Classification":"cgroups",
            "Properties":{
                "root":"/spark-rapids-cgroup",
                "yarn-hierarchy":"yarn"
            }
        }
    ]
```

```
}
```

Amazon EMR 6.x

Exemple de configuration YARN pour Amazon EMR 6.x

```
{
    "Classification":"yarn-site",
    "Properties":{
        "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
        "yarn.resource-types":"yarn.io/gpu",
        "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
        "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/usr/
bin",
        "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
        "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/sys/fs/
cgroup",
        "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
        "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
    }
},{
    "Classification":"container-executor",
    "Properties":{

    },
    "Configurations":[
        {
            "Classification":"gpu",
            "Properties":{
                "module.enabled":"true"
            }
        },
        {
            "Classification":"cgroups",
            "Properties":{
                "root":"/sys/fs/cgroup",
                "yarn-hierarchy":"yarn"
            }
        }
    ]
}
```

3. Configurer Spark pour utiliser RAPIDS

Voici les configurations requises pour permettre à Spark d'utiliser le plug-in RAPIDS :

```
{
 "Classification":"spark-defaults",
 "Properties":{
  "spark.plugins":"com.nvidia.spark.SQLPlugin",
  "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/gpu/
getGpusResources.sh",
  "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-linux/lib:/usr/
local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/cuda/lib:/usr/
local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/native:/docker/usr/
lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native"
 }
}
```

XGBoost4La bibliothèque J-Spark dans XGBoost la documentation est également disponible lorsque le plugin Spark RAPIDS est activé sur votre cluster. Vous pouvez utiliser la configuration suivante pour intégrer XGBoost votre tâche Spark :

```
{
 "Classification":"spark-defaults",
 "Properties":{
  "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-spark_3.0-1.4.2-0.3.0.jar"
 }
}
```

Pour des configurations Spark supplémentaires que vous pouvez utiliser pour régler un cluster EMR accéléré par GPU, reportez-vous au Guide de réglage de Rapids Accelerator pour Apache Spark dans la documentation Nvidia.github.io.

4. Configurer le planificateur de capacité YARN

`DominantResourceCalculator` doit être configuré pour permettre la planification et l'isolation du GPU. Pour plus d'informations, reportez-vous à la section Utilisation du GPU sur YARN dans la documentation d'Apache Hadoop.

```
{
 "Classification":"capacity-scheduler",
 "Properties":{
```

```
  "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
 }
}
```

5. Création d'un fichier JSON pour inclure toutes vos configurations

Vous pouvez créer un fichier JSON contenant votre configuration pour utiliser le plug-in RAPIDS pour votre cluster Spark. Vous fournirez le fichier ultérieurement lors du lancement de votre cluster.

Le fichier peut être stocké localement ou sur S3. Pour plus d'informations sur la manière de fournir des configurations d'applications pour vos clusters, consultez Configuration des applications.

Utilisez les exemples de fichiers suivants comme modèles pour créer vos propres configurations.

Amazon EMR 7.x

Exemple de fichier **`my-configurations.json`** pour Amazon EMR 7.x

```
[
    {
        "Classification":"spark",
        "Properties":{
            "enableSparkRapids":"true"
        }
    },
    {
        "Classification":"yarn-site",
        "Properties":{
            "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
            "yarn.resource-types":"yarn.io/gpu",
            "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
            "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/
usr/bin",
            "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
            "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/spark-
rapids-cgroup",
            "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
            "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
        }
    },
    {
        "Classification":"container-executor",
```

```
        "Properties":{

        },
        "Configurations":[
            {
                "Classification":"gpu",
                "Properties":{
                    "module.enabled":"true"
                }
            },
            {
                "Classification":"cgroups",
                "Properties":{
                    "root":"/spark-rapids-cgroup",
                    "yarn-hierarchy":"yarn"
                }
            }
        ]
    },
    {
        "Classification":"spark-defaults",
        "Properties":{
            "spark.plugins":"com.nvidia.spark.SQLPlugin",
            "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/
gpu/getGpusResources.sh",
            "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-
linux/lib:/usr/local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/
cuda/lib:/usr/local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/
native:/docker/usr/lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native",
            "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-
spark_3.0-1.4.2-0.3.0.jar",
            "spark.rapids.sql.concurrentGpuTasks":"1",
            "spark.executor.resource.gpu.amount":"1",
            "spark.executor.cores":"2",
            "spark.task.cpus":"1",
            "spark.task.resource.gpu.amount":"0.5",
            "spark.rapids.memory.pinnedPool.size":"0",
            "spark.executor.memoryOverhead":"2G",
            "spark.locality.wait":"0s",
            "spark.sql.shuffle.partitions":"200",
            "spark.sql.files.maxPartitionBytes":"512m"
        }
    },
    {
```

```
        "Classification":"capacity-scheduler",
        "Properties":{
            "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
        }
    }
]
```

Amazon EMR 6.x

Exemple de fichier **`my-configurations.json`** pour Amazon EMR 6.x

```
[
    {
        "Classification":"spark",
        "Properties":{
            "enableSparkRapids":"true"
        }
    },
    {
        "Classification":"yarn-site",
        "Properties":{
            "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
            "yarn.resource-types":"yarn.io/gpu",
            "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
            "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/
usr/bin",
            "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
            "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/sys/fs/
cgroup",
            "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
            "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
        }
    },
    {
        "Classification":"container-executor",
        "Properties":{

        },
        "Configurations":[
            {
                "Classification":"gpu",
                "Properties":{
```

```
                    "module.enabled":"true"
                }
            },
            {
                "Classification":"cgroups",
                "Properties":{
                    "root":"/sys/fs/cgroup",
                    "yarn-hierarchy":"yarn"
                }
            }
        ]
    },
    {
        "Classification":"spark-defaults",
        "Properties":{
            "spark.plugins":"com.nvidia.spark.SQLPlugin",
            "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/
gpu/getGpusResources.sh",
            "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-
linux/lib:/usr/local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/
cuda/lib:/usr/local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/
native:/docker/usr/lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native",
            "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-
spark_3.0-1.4.2-0.3.0.jar",
            "spark.rapids.sql.concurrentGpuTasks":"1",
            "spark.executor.resource.gpu.amount":"1",
            "spark.executor.cores":"2",
            "spark.task.cpus":"1",
            "spark.task.resource.gpu.amount":"0.5",
            "spark.rapids.memory.pinnedPool.size":"0",
            "spark.executor.memoryOverhead":"2G",
            "spark.locality.wait":"0s",
            "spark.sql.shuffle.partitions":"200",
            "spark.sql.files.maxPartitionBytes":"512m"
        }
    },
    {
        "Classification":"capacity-scheduler",
        "Properties":{
            "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
        }
    }
```

```
]
```

Ajoutez une action d'amorçage pour votre cluster

Pour plus d'informations sur la façon de fournir des scripts d'action d'amorçage lors de la création d'un cluster, voir la rubrique Bases de l'action d'amorçage du Guide de gestion d'Amazon EMR.

Les exemples de scripts suivants montrent comment créer un fichier d'action bootstrap pour Amazon EMR 6.x et 7.x :

Amazon EMR 7.x

Exemple de fichier **`my-bootstrap-action.sh`** pour Amazon EMR 7.x

Pour utiliser YARN afin de gérer les ressources GPU avec Amazon EMR 7.x, vous devez monter CGroup v1 manuellement sur votre cluster. Vous pouvez le faire à l'aide d'un script d'action d'amorçage, comme indiqué dans cet exemple.

```
#!/bin/bash
set -ex

sudo mkdir -p /spark-rapids-cgroup/devices
sudo mount -t cgroup -o devices cgroupv1-devices /spark-rapids-cgroup/devices
sudo chmod a+rwx -R /spark-rapids-cgroup
```

Amazon EMR 6.x

Exemple de fichier **`my-bootstrap-action.sh`** pour Amazon EMR 6.x

Pour Amazon EMR 6.x, vous devez ouvrir des autorisations CGroup pour YARN sur votre cluster. Pour ce faire, vous pouvez utiliser un script d'action d'amorçage, comme illustré dans cet exemple.

```
#!/bin/bash
set -ex

sudo chmod a+rwx -R /sys/fs/cgroup/cpu,cpuacct
sudo chmod a+rwx -R /sys/fs/cgroup/devices
```

Lancement de votre cluster

La dernière étape consiste à lancer votre cluster avec les configurations de cluster mentionnées ci-dessus. Voici un exemple de commande pour lancer un cluster via la CLI Amazon EMR :

```
 aws emr create-cluster \
--release-label emr-7.9.0 \
--applications Name=Hadoop Name=Spark \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=my-key-pair,InstanceProfile=EMR_EC2_DefaultRole \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=m4.4xlarge \

                  InstanceGroupType=CORE,InstanceCount=1,InstanceType=g4dn.2xlarge \

                  InstanceGroupType=TASK,InstanceCount=1,InstanceType=g4dn.2xlarge \
--configurations file:///my-configurations.json \
--bootstrap-actions Name='My Spark Rapids Bootstrap action',Path=s3://amzn-s3-demo-
bucket/my-bootstrap-action.sh
```

Accès au shell de Spark

Le shell Spark est basé sur le Scala REPL (Read-Eval-Print-Loop). Il vous permet de créer des programmes Spark de façon interactive et de soumettre du travail à l'infrastructure. Vous pouvez accéder au shell de Spark en vous connectant au nœud primaire avec SSH et en invoquant `spark-shell`. Pour plus d'informations sur la connexion au nœud primaire, consultez la section Connexion au nœud primaire à l'aide de SSH dans le Guide de gestion Amazon EMR. Les exemples suivants utilisent les journaux d'accès du serveur HTTP Apache stockés dans Amazon S3.

> **Note**
>
> Le compartiment utilisé dans ces exemples est disponible pour les clients ayant accès à la région USA Est (Virginie du Nord).

Par défaut, le shell Spark crée son propre SparkContextobjet appelé`sc`. Vous pouvez utiliser ce contexte s'il est requis dans le REPL. sqlContextest également disponible en coque et il s'agit d'un HiveContext.

Example Utilisation du shell de Spark pour compter les occurrences d'une chaîne dans un fichier stocké dans Amazon S3

Cet exemple utilise `sc` pour lire un fichier texte stocké dans Amazon S3.

```
scala> sc
res0: org.apache.spark.SparkContext = org.apache.spark.SparkContext@404721db

scala> val textFile = sc.textFile("s3://elasticmapreduce/samples/hive-ads/tables/
impressions/dt=2009-04-13-08-05/ec2-0-51-75-39.amazon.com-2009-04-13-08-05.log")
```

Spark crée le fichier texte et la structure de données associée. Ensuite, l'exemple compte le nombre de lignes figurant dans le fichier journal avec la chaîne « cartoonnetwork.com » :

```
scala> val linesWithCartoonNetwork = textFile.filter(line =>
 line.contains("cartoonnetwork.com")).count()
linesWithCartoonNetwork: org.apache.spark.rdd.RDD[String] = MapPartitionsRDD[2] at
 filter at <console>:23
<snip>
<Spark program runs>
scala> linesWithCartoonNetwork
res2: Long = 9
```

Example Utilisation du shell de Spark basé sur Python pour compter les occurrences d'une chaîne dans un fichier stocké dans Amazon S3

Spark inclut également un shell basé sur Python, `pyspark`, que vous pouvez utiliser pour créer un prototype de programmes Spark écrits en Python. Tout comme avec`spark-shell`, invoke `pyspark` sur le nœud principal ; il possède également le même SparkContextobjet.

```
>>> sc
<pyspark.context.SparkContext object at 0x7fe7e659fa50>
>>> textfile = sc.textFile("s3://elasticmapreduce/samples/hive-ads/tables/impressions/
dt=2009-04-13-08-05/ec2-0-51-75-39.amazon.com-2009-04-13-08-05.log")
```

Spark crée le fichier texte et la structure de données associée. Ensuite, l'exemple compte le nombre de lignes figurant dans le fichier journal avec la chaîne « cartoonnetwork.com ».

```
>>> linesWithCartoonNetwork = textfile.filter(lambda line: "cartoonnetwork.com" in
 line).count()
```

```
15/06/04 17:12:22 INFO lzo.GPLNativeCodeLoader: Loaded native gpl library from the
 embedded binaries
15/06/04 17:12:22 INFO lzo.LzoCodec: Successfully loaded & initialized native-lzo
 library [hadoop-lzo rev EXAMPLE]
15/06/04 17:12:23 INFO fs.EmrFileSystem: Consistency disabled, using
 com.amazon.ws.emr.hadoop.fs.s3n.S3NativeFileSystem as filesystem implementation
<snip>
<Spark program continues>
>>> linesWithCartoonNetwork
9
```

Utiliser Amazon SageMaker Spark pour le machine learning

Lorsque vous utilisez les versions 5.11.0 et ultérieures d'Amazon EMR, le composant `aws-sagemaker-spark-sdk` est installé parallèlement à Spark. Ce composant installe Amazon SageMaker Spark et les dépendances associées pour l'intégration de Spark à Amazon SageMaker. Vous pouvez utiliser Amazon SageMaker Spark pour créer des pipelines d'apprentissage automatique (ML) Spark à l'aide d'Amazon SageMaker Stages. Pour plus d'informations, consultez le fichier README d'Amazon SageMaker Spark sur GitHub et son utilisation d'Apache Spark avec Amazon SageMaker dans le manuel Amazon SageMaker Developer Guide.

Ecriture d'une application Spark

Les applications Spark peuvent être écrites en Scala, Java ou Python. Il existe plusieurs exemples d'applications Spark proposés dans la rubrique Exemples Spark de la documentation Apache Spark. L'exemple d'estimation du nombre Pi est présenté ci-dessous dans les trois applications prises en charge en mode natif. Vous pouvez également consulter des exemples complets dans `$SPARK_HOME/examples` et sur GitHub. Pour plus d'informations sur la façon de compiler JARs pour Spark, consultez la rubrique Démarrage rapide dans la documentation d'Apache Spark.

Scala

Pour éviter les problèmes de compatibilité avec Scala, nous vous conseillons d'utiliser les dépendances Spark pour obtenir la bonne version de Scala lorsque vous compilez une application Spark pour un cluster Amazon EMR. La version de Scala que vous devez utiliser dépend de la version de Spark installée sur votre cluster. Par exemple, la version 5.30.1 d'Amazon EMR utilise Spark 2.4.5, qui est construit avec Scala 2.11. Si votre cluster utilise la version 5.30.1 d'Amazon EMR, utilisez les dépendances Spark pour Scala 2.11. Pour plus d'informations sur les versions de Scala utilisées par Spark, consultez la documentation d'Apache Spark.

```
package org.apache.spark.examples
import scala.math.random
import org.apache.spark._

/** Computes an approximation to pi */
object SparkPi {
  def main(args: Array[String]) {
    val conf = new SparkConf().setAppName("Spark Pi")
    val spark = new SparkContext(conf)
    val slices = if (args.length > 0) args(0).toInt else 2
    val n = math.min(100000L * slices, Int.MaxValue).toInt // avoid overflow
    val count = spark.parallelize(1 until n, slices).map { i =>
      val x = random * 2 - 1
      val y = random * 2 - 1
      if (x*x + y*y < 1) 1 else 0
    }.reduce(_ + _)
    println("Pi is roughly " + 4.0 * count / n)
    spark.stop()
  }
}
```

Java

```
package org.apache.spark.examples;

import org.apache.spark.SparkConf;
import org.apache.spark.api.java.JavaRDD;
import org.apache.spark.api.java.JavaSparkContext;
import org.apache.spark.api.java.function.Function;
import org.apache.spark.api.java.function.Function2;

import java.util.ArrayList;
import java.util.List;

/**
 * Computes an approximation to pi
 * Usage: JavaSparkPi [slices]
 */
public final class JavaSparkPi {

  public static void main(String[] args) throws Exception {
    SparkConf sparkConf = new SparkConf().setAppName("JavaSparkPi");
```

```
    JavaSparkContext jsc = new JavaSparkContext(sparkConf);

    int slices = (args.length == 1) ? Integer.parseInt(args[0]) : 2;
    int n = 100000 * slices;
    List<Integer> l = new ArrayList<Integer>(n);
    for (int i = 0; i < n; i++) {
      l.add(i);
    }

    JavaRDD<Integer> dataSet = jsc.parallelize(l, slices);

    int count = dataSet.map(new Function<Integer, Integer>() {
      @Override
      public Integer call(Integer integer) {
        double x = Math.random() * 2 - 1;
        double y = Math.random() * 2 - 1;
        return (x * x + y * y < 1) ? 1 : 0;
      }
    }).reduce(new Function2<Integer, Integer, Integer>() {
      @Override
      public Integer call(Integer integer, Integer integer2) {
        return integer + integer2;
      }
    });

    System.out.println("Pi is roughly " + 4.0 * count / n);

    jsc.stop();
  }
}
```

Python

```
import argparse
import logging
from operator import add
from random import random

from pyspark.sql import SparkSession

logger = logging.getLogger(__name__)
logging.basicConfig(level=logging.INFO, format="%(levelname)s: %(message)s")
```

```
def calculate_pi(partitions, output_uri):
    """
    Calculates pi by testing a large number of random numbers against a unit circle
    inscribed inside a square. The trials are partitioned so they can be run in
    parallel on cluster instances.

    :param partitions: The number of partitions to use for the calculation.
    :param output_uri: The URI where the output is written, typically an Amazon S3
                       bucket, such as 's3://example-bucket/pi-calc'.
    """

    def calculate_hit(_):
        x = random() * 2 - 1
        y = random() * 2 - 1
        return 1 if x**2 + y**2 < 1 else 0

    tries = 100000 * partitions
    logger.info(
        "Calculating pi with a total of %s tries in %s partitions.", tries, partitions
    )
    with SparkSession.builder.appName("My PyPi").getOrCreate() as spark:
        hits = (
            spark.sparkContext.parallelize(range(tries), partitions)
            .map(calculate_hit)
            .reduce(add)
        )
        pi = 4.0 * hits / tries
        logger.info("%s tries and %s hits gives pi estimate of %s.", tries, hits, pi)
        if output_uri is not None:
            df = spark.createDataFrame([(tries, hits, pi)], ["tries", "hits", "pi"])
            df.write.mode("overwrite").json(output_uri)


if __name__ == "__main__":
    parser = argparse.ArgumentParser()
    parser.add_argument(
        "--partitions",
        default=2,
        type=int,
        help="The number of parallel partitions to use when calculating pi.",
    )
    parser.add_argument(
```

```
        "--output_uri", help="The URI where output is saved, typically an S3 bucket."
    )
    args = parser.parse_args()

    calculate_pi(args.partitions, args.output_uri)
```

Améliorer les performances de Spark avec Amazon S3

Amazon EMR propose des fonctionnalités pour aider à optimiser la performance lors de l'utilisation de Spark pour demander, lire et écrire des données enregistrées dans Amazon S3.

S3 Select peut améliorer la performance de requête pour les fichiers CSV et JSON dans certaines applications en « poussant vers le bas » le traitement à Amazon S3.

Le committer optimisé pour EMRFS S3 est une alternative à la OutputCommitterclasse, qui utilise la fonctionnalité de téléchargement en plusieurs parties d'EMRFS pour améliorer les performances lors de l'écriture de fichiers Parquet sur Amazon S3 à l'aide de Spark et de Datasets. DataFrames

Rubriques

- Utilisation de S3 Select avec Spark pour améliorer les performances des requêtes
- Utilisation d'un valideur EMRFS optimisé pour S3
- Utiliser le protocole de validation optimisé pour EMRFS S3
- Réessayer de demander Amazon S3 avec EMRFS

Utilisation de S3 Select avec Spark pour améliorer les performances des requêtes

> **Important**
>
> Amazon S3 Select n’est plus disponible pour les nouveaux clients. Les clients existants d’Amazon S3 Select peuvent continuer à utiliser cette fonctionnalité comme d’habitude. En savoir plus

À partir de la version 5.17.0 d'Amazon EMR, vous pouvez utiliser S3 Select avec Spark sur Amazon EMR. S3 Select permet aux applications de ne récupérer qu'un sous-ensemble de données d'un objet. Pour Amazon EMR, le travail de calcul qu'implique le filtrage d'importants jeux de données pour le traitement est « poussé » du cluster vers Amazon S3, ce qui peut améliorer les performances de certaines applications et réduire le volume de données transféré entre Amazon EMR et Amazon S3.

S3 Select est pris en charge avec les fichiers CSV et JSON à l'aide des valeurs `s3selectCSV` et `s3selectJSON` pour spécifier le format de données. Pour plus d'informations et d'exemples, consultez Spécifier S3 Select dans votre code.

S3 Select est-il adapté à mon application ?

Nous vous recommandons de tester vos applications avec et sans S3 Select pour voir si son utilisation peut être adaptée à votre application.

Utilisez les consignes suivantes pour déterminer si votre application peut utiliser S3 Select :

- Votre requête exclut plus de la moitié de l'ensemble de données d'origine.
- Votre connexion réseau entre Amazon S3 et le cluster Amazon EMR a une bonne vitesse de transfert et une bande passante disponible. Amazon S3 ne compresse pas les réponses HTTP, de sorte que la taille de la réponse est susceptible d'augmenter pour les fichiers d'entrée compressés.

Considérations et restrictions

- Le chiffrement côté serveur d'Amazon S3 avec des clés de chiffrement fournies par le client (SSE-C) et le chiffrement côté client ne sont pas pris en charge.
- La propriété `AllowQuotedRecordDelimiters` n'est pas prise en charge. Si cette propriété est spécifiée, la requête échoue.
- Seuls les fichiers CSV et JSON au format UTF-8 sont pris en charge. CSVsLes lignes multiples ne sont pas prises en charge.
- Seuls les fichiers non compressés ou gzip sont pris en charge.
- Les options Spark CSV et JSON telles que `nanValue`, `positiveInf`, `negativeInf` et les options liées aux enregistrements corrompus (par exemple, les modes failfast et dropmalformed) ne sont pas prises en charge.
- L'utilisation de virgules (,) avec des décimales n'est pas prise en charge. Par exemple, `10,000` n'est pas pris en charge, mais `10000` l'est.

- Les caractères de commentaire ne sont pas pris en charge dans la dernière ligne.
- Des lignes vides à la fin d'un fichier ne sont pas traitées.
- Les filtres suivants ne sont pas transmis à Amazon S3 :
  - Fonctions d'agrégation, telles que `COUNT()` et `SUM()`.
  - Filtres qui `CAST()` un attribut. Par exemple, `CAST(stringColumn as INT) = 1`.
  - Filtres avec un attribut d'objet ou complexe. Par exemple, `intArray[1] = 1`, `objectColumn.objectNumber = 1`.
  - Filtres pour lesquels la valeur n'est pas une valeur littérale. Par exemple, `intColumn1 = intColumn2`
  - Seuls les types de données S3 Select pris en charge sont pris en charge par les limitations de documentation.

Spécifier S3 Select dans votre code

Les exemples suivants montrent comment spécifier S3 Select pour CSV à l'aide de Scala, SQL, R et PySpark. Vous pouvez utiliser S3 Select pour JSON de la même manière. Pour une liste d'options, leurs valeurs par défaut et leurs limitations, consultez Options.

PySpark

```
spark
  .read
  .format("s3selectCSV") // "s3selectJson" for Json
  .schema(...) // optional, but recommended
  .options(...) // optional
  .load("s3://path/to/my/datafiles")
```

R

```
read.df("s3://path/to/my/datafiles", "s3selectCSV", schema, header = "true",
 delimiter = "\t")
```

Scala

```
spark
  .read
  .format("s3selectCSV") // "s3selectJson" for Json
```

```
  .schema(...) // optional, but recommended
  .options(...) // optional. Examples:
  // .options(Map("quote" -> "\'", "header" -> "true")) or
  // .option("quote", "\'").option("header", "true")
  .load("s3://path/to/my/datafiles")
```

SQL

```
CREATE TEMPORARY VIEW MyView (number INT, name STRING) USING s3selectCSV OPTIONS
 (path "s3://path/to/my/datafiles", header "true", delimiter "\t")
```

Options

Les options suivantes sont disponibles lors de l'utilisation de `s3selectCSV` et `s3selectJSON`. Si rien n'est spécifié, les valeurs par défaut sont utilisées.

Options avec S3selectCSV

Option	Par défaut	Utilisation
`compression`	`"none"`	Indique si la compression est utilisée. `"gzip"` est le seul paramètre pris en charge, à part `"none"`.
`delimiter`	`","`	Spécifie le caractère délimiteur de champ.
`quote`	`'\"'`	Spécifie le caractère de citation. La spécification d'une chaîne vide n'est pas prise en charge et mène à une erreur de XML malformé.
`escape`	`'\\'`	Spécifie le caractère échappe.
`header`	`"false"`	`"false"` spécifie qu'il n'y a pas d'en-tête. `"true"` spécifie qu'une en-tête est dans la

Option	Par défaut	Utilisation
		première ligne. Les en-têtes ne sont prises en charge que pour la première ligne, et les lignes vides avant une en-tête ne sont pas prises en charge.
comment	`"#"`	Spécifie le caractère de commentaire. L'indicateur de commentaire ne peut pas être désactivé. En d'autres termes, la valeur de `\u0000` n'est pas prise en charge.
`nullValue`	`""`	

Options avec S3selectJSON

Option	Par défaut	Utilisation
`compression`	`"none"`	Indique si la compression est utilisée. `"gzip"` est le seul paramètre pris en charge, à part `"none"`.
`multiline`	« false »	`"false"` spécifie que JSON est en format `LINES` S3 Select. Cela implique que chaque ligne des données d'entrée contient un seul objet JSON. `"true"` spécifie que JSON est en format `DOCUMENT` S3 Select, ce qui signifie qu'un objet JSON peut s'étendre sur plusieurs lignes dans les données d'entrée.

Utilisation d'un valideur EMRFS optimisé pour S3

Le committer optimisé pour EMRFS S3 est une OutputCommitterimplémentation alternative optimisée pour écrire des fichiers sur Amazon S3 lors de l'utilisation d'EMRFS. Le validateur EMRFS optimisé pour S3 améliore les performances de l'application en évitant les opérations de liste et de renommage effectuées dans Amazon S3 pendant les phases de validation des tâches. Le valideur est disponible dans les versions 5.19.0 et ultérieures d'Amazon EMR et est activé par défaut dans les versions 5.20.0 et ultérieures d'Amazon EMR. Le committer est utilisé pour les tâches Spark qui utilisent Spark ou Datasets. DataFrames À partir d'Amazon EMR 6.4.0, ce validateur peut être utilisé pour tous les formats courants, notamment parquet, ORC et les formats basés sur le texte (notamment CSV et JSON). Pour les versions antérieures à Amazon EMR 6.4.0, seul le format Parquet est pris en charge. Dans certaines situations, le valideur n'est pas utilisé. Pour de plus amples informations, veuillez consulter Exigences du valideur EMRFS optimisé pour S3.

Rubriques

- Exigences du valideur EMRFS optimisé pour S3
- Le valideur EMRFS optimisé pour S3 et les chargements partitionnés
- Considérations de réglage de tâche
- Activer le validateur EMRFS S3 pour Amazon EMR 5.19.0

Exigences du valideur EMRFS optimisé pour S3

Le valideur EMRFS optimisé pour S3 est utilisé lorsque les conditions suivantes sont réunies :

- Vous exécutez des tâches Spark qui utilisent Spark ou Datasets pour écrire des fichiers sur Amazon S3. DataFrames À partir d'Amazon EMR 6.4.0, ce validateur peut être utilisé pour tous les formats courants, notamment parquet, ORC et les formats basés sur le texte (notamment CSV et JSON). Pour les versions antérieures à Amazon EMR 6.4.0, seul le format Parquet est pris en charge.
- Les chargements partitionnés sont activés dans Amazon EMR. Il s’agit de l’option par défaut. Pour de plus amples informations, veuillez consulter Le valideur EMRFS optimisé pour S3 et les chargements partitionnés.
- Le support de format de fichier intégré à Spark est utilisé. La prise en charge du format de fichier intégré est utilisée dans les circonstances suivantes :
  - Pour les tables de métastore Hive, lorsque `spark.sql.hive.convertMetastoreParquet` est défini sur `true` pour les tables Parquet, ou lorsque

`spark.sql.hive.convertMetastoreOrc` est défini sur `true` pour les tables Orc avec Amazon EMR 6.4.0 ou version ultérieure. Il s'agit des paramètres par défaut.

- Lorsque les tâches écrivent dans des sources de données ou des tables au format de fichier – par exemple, la table cible est créée avec la clause `USING parquet`.
- Lorsque des tâches écrivent dans des tableaux Parquet de metastore Hive non partitionnées. Le support Parquet intégré à Spark ne prend pas en charge avec les tableaux Hive partitionnés. Il s'agit d'une limitation connue. Pour plus d'informations, consultez la section Conversion des tables Hive Metastore Parquet dans le guide Apache Spark DataFrames et Datasets.

- Les opérations de tâche Spark qui écrivent dans un emplacement de partition par défaut (par exemple, `${table_location}/k1=v1/k2=v2/`) utilisent le valideur. Le valideur n'est pas utilisé si une opération de tâche écrit dans un emplacement de partition personnalisé, par exemple, si un emplacement de partition personnalisé est défini à l'aide de la commande `ALTER TABLE SQL`.
- Les valeurs suivantes pour Spark doivent être utilisées :
  - La propriété `spark.sql.parquet.fs.optimized.committer.optimization-enabled` doit être définie sur `true`. Il s'agit du paramètre par défaut avec Amazon EMR 5.20.0 ou version suivante. Avec Amazon EMR 5.19.0, la valeur par défaut est `false`. Pour plus d'informations sur la configuration de cette valeur, consultez Activer le validateur EMRFS S3 pour Amazon EMR 5.19.0.
  - Si vous écrivez dans des tables de métastore Hive non partitionnées, seuls les formats de fichier Parquet et Orc sont pris en charge. `spark.sql.hive.convertMetastoreParquet`doit être défini sur `true` si vous écrivez sur des tables de métastore Parquet Hive non partitionnées. `spark.sql.hive.convertMetastoreOrc`doit être défini sur `true` si vous écrivez dans des tables de métastore Orc Hive non partitionnées. Il s'agit des paramètres par défaut.
  - `spark.sql.parquet.output.committer.class` doit être défini sur `com.amazon.emr.committer.EmrOptimizedSparkSqlParquetOutputCommitter`. Il s’agit du paramètre par défaut.
  - `spark.sql.sources.commitProtocolClass` doit être défini sur `org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol` ou `org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol`. `org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol` est le paramètre par défaut pour les versions 5.30.0 et ultérieures de la série Amazon EMR 5.x, et pour les versions 6.2.0 et ultérieures de la série Amazon EMR 6.x. `org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol` est le paramètre par défaut pour les versions précédentes d’Amazon EMR.

- Si des tâches Spark remplacent des ensembles de données Parquet partitionnés par des colonnes de partition dynamique, l'option d'écriture `partitionOverwriteMode` et `spark.sql.sources.partitionOverwriteMode` doivent être définis sur `static`. Il s'agit du paramètre par défaut.

> **Note**
>
> L'option d'écriture `partitionOverwriteMode` a été introduite dans Spark 2.4.0. Pour Spark version 2.3.2, inclus avec Amazon EMR version 5.19.0, définissez la propriété `spark.sql.sources.partitionOverwriteMode`.

Cas dans lesquels le valideur EMRFS optimisé pour S3 n'est pas utilisé

En général, le valideur EMRFS S3 optimisé n'est pas utilisé dans les situations suivantes.

Situation	Pourquoi le validateur n'est pas utilisé
Lorsque vous écrivez dans HDFS	Le validateur prend uniquement en charge l'écriture sur Amazon S3 à l'aide d'EMRFS.
Lorsque vous utilisez le système de fichiers S3A	Le validateur ne prend en charge que EMRFS.
Lorsque vous utilisez l' MapReduce API RDD de Spark	Le committer prend uniquement en charge l'utilisation de SparkSQL ou DataFrame Dataset. APIs

Les exemples en Scala suivants illustrent quelques situations supplémentaires qui empêchent d'utiliser le validateur optimisé pour EMRFS S3 en totalité (premier exemple) et en partie (deuxième exemple).

Example – Mode de remplacement de partition dynamique

L'exemple Scala suivant indique à Spark d'utiliser un algorithme de validation différent, ce qui empêche complètement l'utilisation du validateur optimisé pour EMRFS S3. Le code définit la propriété `partitionOverwriteMode` sur `dynamic` pour n'écraser que les partitions sur lesquelles

vous écrivez des données. Ensuite, les colonnes de partition dynamique sont spécifiées par `partitionBy`, et le mode d'écriture est défini sur `overwrite`.

```
val dataset = spark.range(0, 10)
  .withColumn("dt", expr("date_sub(current_date(), id)"))

dataset.write.mode("overwrite")
  .option("partitionOverwriteMode", "dynamic")
  .partitionBy("dt")
  .parquet("s3://amzn-s3-demo-bucket1/output")
```

Vous devez configurer les trois paramètres pour éviter d'utiliser le validateur EMRFS S3. Lorsque vous le faites, Spark exécute un algorithme de validation différent qui est spécifié dans le protocole de validation de Spark. Pour les versions d'Amazon EMR 5.x antérieures à la version 5.30.0 et pour les versions d'Amazon EMR 6.x antérieures à la version 6.2.0, le protocole de validation utilise le répertoire de préparation de Spark, qui est un répertoire temporaire créé sous l'emplacement de sortie commençant par `.spark-staging`. L'algorithme renomme séquentiellement les répertoires de partition, ce qui peut avoir un impact négatif sur les performances. Pour de plus amples informations sur Amazon EMR versions 5.30.0 et versions ultérieures, ainsi que 6.2.0 et versions ultérieures, veuillez consulter Utiliser le protocole de validation optimisé pour EMRFS S3.

L'algorithme dans Spark 2.4.0 exécute les étapes suivantes :

1. Les tentatives de tâches écrivent leur résultat dans les répertoires de partition situés dans le répertoire de préparation de Spark, par exemple, `${outputLocation}/spark-staging-${jobID}/k1=v1/k2=v2/`.
2. Pour chaque partition écrite, la tâche tente de suivre les chemins de partition relatifs, par exemple, `k1=v1/k2=v2`.
3. Lorsqu'une tâche se termine avec succès elle fournit le pilote avec tous les chemins de partition relatifs qu'elle a suivis.
4. Une fois toutes les tâches terminées, la phase de validation de tâche collecte tous les répertoires de partition que les tentatives de tâche réussies ont écrit dans le répertoire intermédiaire de Spark. Spark renomme de manière séquentielle chacun de ces répertoires dans son emplacement de sortie final à l'aide d'opérations attribution de nouveau nom de l'arborescence.
5. Le répertoire intermédiaire est supprimé avant que la phase de validation de tâche soit terminée.

Example – Emplacement de partition personnalisé

Dans cet exemple, le code Scala s'insère dans deux partitions. Une partition possède un emplacement de partition personnalisé. L'autre partition utilise l'emplacement de partition par défaut. Le valideur EMRFS optimisé pour S3 est utilisé uniquement pour écrire la sortie de tâche dans la partition qui utilise l'emplacement de partition par défaut.

```
val table = "dataset"
val location = "s3://bucket/table"

spark.sql(s"""
  CREATE TABLE $table (id bigint, dt date)
  USING PARQUET PARTITIONED BY (dt)
  LOCATION '$location'
""")

// Add a partition using a custom location
val customPartitionLocation = "s3://bucket/custom"
spark.sql(s"""
  ALTER TABLE $table ADD PARTITION (dt='2019-01-28')
  LOCATION '$customPartitionLocation'
""")

// Add another partition using default location
spark.sql(s"ALTER TABLE $table ADD PARTITION (dt='2019-01-29')")

def asDate(text: String) = lit(text).cast("date")

spark.range(0, 10)
  .withColumn("dt",
    when($"id" > 4, asDate("2019-01-28")).otherwise(asDate("2019-01-29")))
  .write.insertInto(table)
```

Le code Scala crée les objets Amazon S3 suivants :

```
custom/part-00001-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
custom_$folder$
table/_SUCCESS
table/dt=2019-01-29/part-00000-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
table/dt=2019-01-29_$folder$
table_$folder$
```

Lorsque vous écrivez des partitions dans des emplacements personnalisés, Spark utilise un algorithme de validation similaire à celui de l'exemple précédent, qui est décrit ci-dessous. Comme dans l'exemple précédent, l'algorithme se traduit par des attributions séquentielles de nouveaux noms, ce qui peut avoir un impact négatif sur les performances.

1. Lors de l'écriture de la sortie d'une partition dans un emplacement personnalisé, les tâches écrivent un fichier dans le répertoire intermédiaire de Spark, qui est créé sous l'emplacement de sortie final. Le nom du fichier comprend un UUID aléatoire pour éviter les conflits de fichier. La tentative de tâche suit chaque fichier, ainsi que le chemin de la sortie souhaité final.
2. Lorsqu'une tâche se termine avec succès, elle fournit au pilote les fichiers et les chemins de sortie souhaités finaux.
3. Une fois toutes les tâches terminées, la phase de validation de tâche renomme de manière séquentielle tous les fichiers qui ont été écrits pour les partitions dans les emplacements personnalisés en leurs chemins de sortie finaux.
4. Le répertoire intermédiaire est supprimé avant que la phase de validation de tâche soit terminée.

Le valideur EMRFS optimisé pour S3 et les chargements partitionnés

Pour utiliser le valideur EMRFS optimisé pour S3, vous devez activer les chargements partitionnés dans Amazon EMR. Les chargements partitionnés sont activés par défaut. Vous pouvez les réactiver si besoin est. Pour plus d'informations, consultez Configuration d'un chargement partitionné pour Simple Storage Service (Amazon S3) dans le Guide de gestion Amazon EMR.

Le valideur EMRFS optimisé pour S3 utilise des caractéristiques similaires à des transactions de chargements partitionnés pour assurer que les fichiers écrits par des tentatives de tâches apparaissent uniquement dans l'emplacement de sortie du travail lors de la validation de tâche. En utilisant les téléchargements partitionnés de cette manière, le validateur améliore les performances de validation des tâches par rapport à la version 2 de l' FileOutputCommitter algorithme par défaut. Lors de l'utilisation d'un valideur EMRFS optimisé pour S3, il existe certaines différences importantes à prendre en compte comparé au comportement de chargement partitionné :

- Les chargements partitionnés sont toujours effectués, peu importe la taille du fichier. Ceci diffère du comportement normal d'EMRFS, où la propriété `fs.s3n.multipart.uploads.split.size` contrôle la taille du fichier à laquelle les chargements partitionnés sont enclenchés.
- Les chargements partitionnés sont laissés en état non terminé pendant un plus long laps de temps jusqu'à ce que la tâche soit validée ou abandonnée. Ceci diffère du comportement normal

d'EMRFS, où un chargement partitionné se termine lorsqu'un tâche termine d'écrire un fichier donné.

En raison de ces différences, si un Spark Executor JVM crache ou qu'il est tué lorsque des tâches sont en cours d'exécution ou écrivent des données à Amazon S3, les chargements partitionnés inachevés ont plus de chances d'être laissés de côté. Pour cette raison, lorsque vous utilisez un valideur EMRFS optimisé pour S3, assurez-vous de suivre les bonnes pratiques de gestion de chargements partitionnés échoués. Pour plus d'informations, consultez Bonnes pratiques pour travailler avec des compartiments Amazon S3 dans le Guide de gestion Amazon EMR.

Considérations de réglage de tâche

Le valideur EMRFS optimisé pour S3 consomme une petite quantité de mémoire pour chaque fichier écrit par une tentative de tâche, jusqu'à ce que la tâche soit validée ou abandonnée. La quantité de mémoire consommée est négligeable dans la plupart des tâches. Pour les tâches dotées de tâches à longue exécution qui écrivent un grand nombre de fichiers, la mémoire consommée par le valideur peut être significative et nécessitera des ajustements à l'allocation de mémoire aux exécuteurs Spark. Vous pouvez ajuster la mémoire d'exécuteur à l'aide de la propriété `spark.executor.memory`. Une seule tâche écrivant 100 000 fichiers nécessiterait en général environ 100 Mo de mémoire supplémentaire. Pour plus d'informations, consultez Propriétés d'applications dans la Documentation de configuration Apache Spark.

Activer le validateur EMRFS S3 pour Amazon EMR 5.19.0

Si vous utilisez Amazon EMR 5.19.0, vous pouvez définir la propriété `spark.sql.parquet.fs.optimized.committer.optimization-enabled` sur `true` lorsque vous créez un cluster, ou depuis Spark si vous utilisez Amazon EMR.

Activation d'un valideur EMRFS optimisé pour S3 lors de la création d'un cluster

Utilisez la classification de configuration `spark-defaults` pour définir la propriété `spark.sql.parquet.fs.optimized.committer.optimization-enabled` sur `true`. Pour de plus amples informations, veuillez consulter Configuration des applications.

Activation d'un valideur EMRFS optimisé pour S3 depuis Spark

Vous pouvez définir `spark.sql.parquet.fs.optimized.committer.optimization-enabled` sur `true` en l'ancrant dans un `SparkConf`, en le passant comme paramètre `--conf` dans

le shell Spark ou les outils `spark-submit` et `spark-sql`, ou dans `conf/spark-defaults.conf`. Pour plus d'informations, consultez Configuration Spark dans la documentation Apache Spark.

L'exemple suivant montre comment activer le valideur en exécutant une commande spark-sql.

```
spark-sql \
  --conf spark.sql.parquet.fs.optimized.committer.optimization-enabled=true \
  -e "INSERT OVERWRITE TABLE target_table SELECT * FROM source_table;"
```

Utiliser le protocole de validation optimisé pour EMRFS S3

Le protocole de validation optimisé pour EMRFS S3 est une FileCommitProtocolimplémentation alternative optimisée pour écrire des fichiers avec remplacement dynamique de partition Spark sur Amazon S3 lors de l'utilisation d'EMRFS. Le protocole améliore les performances des applications en évitant les opérations de renommage dans Amazon S3 pendant la phase de validation de la tâche de remplacement de partition dynamique Spark.

Notez que Utilisation d'un valideur EMRFS optimisé pour S3 améliore également les performances en évitant les opérations de renommage. Cependant, cela ne fonctionne pas pour les cas de remplacement de partition dynamique, tandis que les améliorations du protocole de validation ne ciblent que les cas de remplacement de partition dynamique.

Le protocole de validation est disponible avec les versions 5.30.0 et suivantes d'Amazon EMR et 6.2.0 et suivantes, et il est activé par défaut. Amazon EMR a ajouté une amélioration du parallélisme à partir de la version 5.31.0. Le protocole est utilisé pour les tâches Spark qui utilisent Spark ou Datasets. DataFrames Dans certaines circonstances, le protocole de validation n'est pas utilisé. Pour de plus amples informations, veuillez consulter Conditions requises pour le protocole de validation EMRFS optimisé pour S3.

Rubriques

- Conditions requises pour le protocole de validation EMRFS optimisé pour S3
- Le protocole de validation EMRFS optimisé pour S3 et les chargements partitionnés
- Considérations de réglage de tâche

Conditions requises pour le protocole de validation EMRFS optimisé pour S3

Le protocole de validation EMRFS optimisé pour S3 est utilisé lorsque les conditions suivantes sont remplies :

- Vous exécutez des tâches Spark qui utilisent Spark ou Datasets pour remplacer les tables partitionnées. DataFrames
- Vous exécutez des tâches Spark dont le mode de remplacement de partition est `dynamic`.
- Les chargements partitionnés sont activés dans Amazon EMR. Il s'agit de l'option par défaut. Pour de plus amples informations, veuillez consulter Le protocole de validation EMRFS optimisé pour S3 et les chargements partitionnés.
- Le cache du système de fichiers pour EMRFS est activé. Il s'agit de l'option par défaut. Vérifiez que le paramètre `fs.s3.impl.disable.cache` est défini sur `false`.
- Le support intégré des sources de données de Spark est utilisé. La prise en charge de la source de données intégrée est utilisée dans les circonstances suivantes :
  - Lorsque les tâches écrivent dans des sources de données ou des tables intégrées.
  - Lorsque les tâches écrivent dans la table Parquet du métastore Hive. Cela se produit lorsque `spark.sql.hive.convertInsertingPartitionedTable` et `spark.sql.hive.convertMetastoreParquet` sont tous deux définis sur true. Il s'agit des paramètres par défaut.
  - Lorsque les jobs écrivent dans la table ORC du métastore Hive. Cela se produit lorsque `spark.sql.hive.convertInsertingPartitionedTable` et `spark.sql.hive.convertMetastoreOrc` sont tous deux définis sur `true`. Il s'agit des paramètres par défaut.
- Les opérations de tâches Spark qui écrivent dans un emplacement de partition par défaut (par exemple, `${table_location}/k1=v1/k2=v2/`) utilisent le protocole de validation. Le protocole n'est pas utilisé si une opération de tâche écrit dans un emplacement de partition personnalisé (par exemple, si un emplacement de partition personnalisé est défini à l'aide de la commande `ALTER TABLE SQL`).
- Les valeurs suivantes pour Spark doivent être utilisées :
  - `spark.sql.sources.commitProtocolClass` doit être défini sur `org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol`. Il s'agit du paramètre par défaut pour les versions 5.30.0 et supérieures d'Amazon EMR, ainsi que pour les versions 6.2.0 et supérieures.
  - L'option d'écriture `partitionOverwriteMode` ou `spark.sql.sources.partitionOverwriteMode` doit être définie sur `dynamic`. Le paramètre par défaut est `static`.

> **Note**
>
> L'option d'écriture `partitionOverwriteMode` a été introduite dans Spark 2.4.0. Pour Spark version 2.3.2, inclus avec Amazon EMR version 5.19.0, définissez la propriété `spark.sql.sources.partitionOverwriteMode`.

- Si les tâches Spark remplacent la table Parquet du métastore Hive, `spark.sql.hive.convertMetastoreParquet`, `spark.sql.hive.convertInsertingPartitionedTable` et `spark.sql.hive.convertMetastore.partitionOverwriteMode` doivent être définis sur `true`. Il s'agit des paramètres par défaut.
- Si les tâches Spark remplacent la table ORC du métastore Hive, `spark.sql.hive.convertMetastoreOrc`, `spark.sql.hive.convertInsertingPartitionedTable` et `spark.sql.hive.convertMetastore.partitionOverwriteMode` doivent être définis sur `true`. Il s'agit des paramètres par défaut.

Example – Mode de remplacement de partition dynamique

Dans cet exemple de Scala, l'optimisation est déclenchée. Tout d'abord, vous définissez la propriété `partitionOverwriteMode` sur `dynamic`. Cela remplace uniquement les partitions sur lesquelles vous écrivez des données. Ensuite, vous spécifiez les colonnes de partition dynamique avec `partitionBy` et définissez le mode d'écriture sur `overwrite`.

```
val dataset = spark.range(0, 10)
  .withColumn("dt", expr("date_sub(current_date(), id)"))

dataset.write.mode("overwrite")                 // "overwrite" instead of "insert"
  .option("partitionOverwriteMode", "dynamic")  // "dynamic" instead of "static"
  .partitionBy("dt")                            // partitioned data instead of
 unpartitioned data
  .parquet("s3://amzn-s3-demo-bucket1/output")    // "s3://" to use Amazon EMR file
 system, instead of "s3a://" or "hdfs://"
```

Lorsque le protocole de validation EMRFS optimisé pour S3 n'est pas utilisé

En général, le protocole de validation optimisé pour EMRFS S3 fonctionne de la même manière que le protocole de validation Spark par défaut open source.

`org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol`
L'optimisation ne se produira pas dans les situations suivantes.

Situation	Pourquoi le protocole de validation n'est pas utilisé
Lorsque vous écrivez dans HDFS	Le protocole de validation prend uniquement en charge l'écriture sur Amazon S3 à l'aide d'EMRFS.
Lorsque vous utilisez le système de fichiers S3A	Le protocole de validation ne prend en charge que le protocole EMRFS.
Lorsque vous utilisez l' MapReduce API RDD de Spark	Le protocole de validation ne prend en charge que l'utilisation de SparkSQL ou de DataFrame Dataset. APIs
Lorsque le remplacement dynamique de la partition n'est pas déclenché	Le protocole de validation optimise uniquement les cas de remplacement dynamique de partitions. Pour les autres cas, consultez Utilisation d'un valideur EMRFS optimisé pour S3.

Les exemples Scala suivants démontrent quelques situations supplémentaires que le protocole de validation EMRFS optimisé pour S3 délègue à `SQLHadoopMapReduceCommitProtocol`.

Example – Mode de remplacement dynamique de partition avec emplacement de partition personnalisé

Dans cet exemple, les programmes Scala remplacent deux partitions en mode de remplacement dynamique de partition. Une partition possède un emplacement de partition personnalisé. L'autre partition utilise l'emplacement de partition par défaut. Le protocole de validation EMRFS optimisé pour S3 n'améliore que la partition qui utilise l'emplacement de partition par défaut.

```
val table = "dataset"
val inputView = "tempView"
val location = "s3://bucket/table"
```

```
spark.sql(s"""
  CREATE TABLE $table (id bigint, dt date)
  USING PARQUET PARTITIONED BY (dt)
  LOCATION '$location'
""")

// Add a partition using a custom location
val customPartitionLocation = "s3://bucket/custom"
spark.sql(s"""
  ALTER TABLE $table ADD PARTITION (dt='2019-01-28')
  LOCATION '$customPartitionLocation'
""")

// Add another partition using default location
spark.sql(s"ALTER TABLE $table ADD PARTITION (dt='2019-01-29')")

def asDate(text: String) = lit(text).cast("date")

spark.range(0, 10)
  .withColumn("dt",
    when($"id" > 4, asDate("2019-01-28")).otherwise(asDate("2019-01-29")))
  .createTempView(inputView)

// Set partition overwrite mode to 'dynamic'
spark.sql(s"SET spark.sql.sources.partitionOverwriteMode=dynamic")

spark.sql(s"INSERT OVERWRITE TABLE $table SELECT * FROM $inputView")
```

Le code Scala crée les objets Amazon S3 suivants :

```
custom/part-00001-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
custom_$folder$
table/_SUCCESS
table/dt=2019-01-29/part-00000-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
table/dt=2019-01-29_$folder$
table_$folder$
```

> **Note**
>
> L'écriture sur des emplacements de partition personnalisés dans les versions antérieures de Spark peut entraîner une perte de données. Dans cet exemple, la partition `dt='2019-01-28'` serait perdue. Pour plus de détails, consultez SPARK-35106. Ce

problème est résolu dans les versions 5.33.0 et ultérieures d'Amazon EMR, à l'exception des versions 6.0.x et 6.1.x.

Lorsque vous écrivez des partitions dans des emplacements personnalisés, Spark utilise un algorithme de validation similaire à celui de l'exemple précédent, qui est décrit ci-dessous. Comme dans l'exemple précédent, l'algorithme se traduit par des attributions séquentielles de nouveaux noms, ce qui peut avoir un impact négatif sur les performances.

L'algorithme dans Spark 2.4.0 exécute les étapes suivantes :

1. Lors de l'écriture de la sortie d'une partition dans un emplacement personnalisé, les tâches écrivent un fichier dans le répertoire intermédiaire de Spark, qui est créé sous l'emplacement de sortie final. Le nom du fichier comprend un UUID aléatoire pour éviter les conflits de fichier. La tentative de tâche suit chaque fichier, ainsi que le chemin de la sortie souhaité final.
2. Lorsqu'une tâche se termine avec succès, elle fournit au pilote les fichiers et les chemins de sortie souhaités finaux.
3. Une fois toutes les tâches terminées, la phase de validation de tâche renomme de manière séquentielle tous les fichiers qui ont été écrits pour les partitions dans les emplacements personnalisés en leurs chemins de sortie finaux.
4. Le répertoire intermédiaire est supprimé avant que la phase de validation de tâche soit terminée.

Le protocole de validation EMRFS optimisé pour S3 et les chargements partitionnés

Pour utiliser l'optimisation pour le remplacement dynamique des partitions dans le protocole de validation optimisé pour EMRFS S3, les chargements partitionnés doivent être activés dans Amazon EMR. Les chargements partitionnés sont activés par défaut. Vous pouvez les réactiver si besoin est. Pour plus d'informations, consultez Configuration d'un chargement partitionné pour Simple Storage Service (Amazon S3) dans le Guide de gestion Amazon EMR.

Pendant l'écrasement dynamique de partition, le protocole de validation EMRFS optimisé pour S3 utilise les caractéristiques de transaction des chargements partitionnés pour s'assurer que les fichiers écrits par les tentatives de tâches n'apparaissent que dans l'emplacement de sortie de la tâche au moment de la validation de la tâche. En utilisant les chargements partitionnés de cette manière, le protocole de validation améliore les performances de validation des tâches par rapport au protocole par défaut `SQLHadoopMapReduceCommitProtocol`. Lors de l'utilisation du protocole de validation

EMRFS optimisé pour S3, il y a quelques différences clés à prendre en compte par rapport au comportement traditionnel des chargements partitionnés :

- Les chargements partitionnés sont toujours effectués, peu importe la taille du fichier. Ceci diffère du comportement normal d'EMRFS, où la propriété `fs.s3n.multipart.uploads.split.size` contrôle la taille du fichier à laquelle les chargements partitionnés sont enclenchés.
- Les chargements partitionnés sont laissés en état non terminé pendant un plus long laps de temps jusqu'à ce que la tâche soit validée ou abandonnée. Ceci diffère du comportement normal d'EMRFS, où un chargement partitionné se termine lorsqu'un tâche termine d'écrire un fichier donné.

En raison de ces différences, si une JVM Spark Executor se bloque ou est mise hors service alors que des tâches sont en cours d'exécution et écrivent des données sur Amazon S3, ou si une JVM Spark Driver se bloque ou est mise hors service alors qu'une tâche est en cours d'exécution, les chargements partitionnés incomplets sont plus susceptibles d'être laissés derrière. Pour cette raison, lorsque vous utilisez le protocole de validation EMRFS optimisé pour S3, assurez-vous de suivre les bonnes pratiques de gestion de chargements partitionnés échoués. Pour plus d'informations, consultez Bonnes pratiques pour travailler avec des compartiments Amazon S3 dans le Guide de gestion Amazon EMR.

Considérations de réglage de tâche

Sur les exécuteurs Spark, le protocole de validation EMRFS optimisé pour S3 consomme une petite quantité de mémoire pour chaque fichier écrit par une tentative de tâche, jusqu'à ce que la tâche soit validée ou abandonnée. La quantité de mémoire consommée est négligeable dans la plupart des tâches.

Sur les pilotes Spark, le protocole de validation optimisé pour EMRFS S3 nécessite de la mémoire pour stocker les informations de métadonnées de chaque fichier validé jusqu'à ce que la tâche soit validée ou abandonnée. Dans la plupart des tâches, le paramètre de mémoire par défaut du pilote Spark est négligeable.

Pour les tâches qui ont des tâches de longue durée qui écrivent un grand nombre de fichiers, la mémoire que le protocole de validation consomme peut être perceptible et nécessiter des ajustements de la mémoire allouée pour Spark, en particulier pour les exécuteurs Spark. Vous pouvez régler la mémoire à l'aide de la propriété `spark.driver.memory` pour les pilotes Spark et de la propriété `spark.executor.memory` pour les exécuteurs Spark. Une seule tâche écrivant 100 000 fichiers nécessiterait en général environ 100 Mo de mémoire supplémentaire. Pour plus

d'informations, consultez Propriétés d'applications dans la Documentation de configuration Apache Spark.

Réessayer de demander Amazon S3 avec EMRFS

Cette rubrique fournit des informations sur les stratégies de nouvelle tentative que vous pouvez utiliser lorsque vous envoyez des demandes à Amazon S3 avec EMRFS. Lorsque votre taux de demandes augmente, S3 essaie de s'adapter au nouveau taux. Au cours de ce processus, S3 peut limiter les demandes et renvoyer une erreur `503 Slow Down`. Pour améliorer le taux de réussite de vos demandes S3, vous pouvez ajuster votre stratégie de réessai en configurant les propriétés dans votre configuration `emrfs-site`.

Vous pouvez ajuster votre stratégie de réessai de différentes manières.

- Augmentez la limite maximale de tentatives pour la stratégie de backoff exponentiel par défaut.
- Activez et configurez la stratégie de réessai additive-augmentation/diminution multiplicative (AIMD). AIMD est pris en charge pour Amazon EMR versions 6.4.0 et ultérieures.

Utiliser la stratégie de backoff exponentiel par défaut

Par défaut, EMRFS utilise une stratégie de backoff exponentiel pour réessayer les requêtes Amazon S3. La limite de tentatives EMRFS par défaut est de 15. Pour éviter une erreur S3 `503 Slow Down`, vous pouvez augmenter le nombre limite de tentatives lorsque vous créez un nouveau cluster, sur un cluster en cours d'exécution ou lors de l'exécution de l'application.

Pour augmenter le nombre limite de tentatives, vous devez modifier la valeur de `fs.s3.maxRetries` dans votre configuration `emrfs-site`. L'exemple de configuration suivant définit `fs.s3.maxRetries` sur une valeur personnalisée de 30.

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "fs.s3.maxRetries": "30"
      }
    }
]
```

Pour plus d'informations sur l'utilisation des objets de configuration, consultez Configuration des applications.

Utiliser la stratégie de nouvelle tentative de l'AIMD

Avec les versions 6.4.0 et ultérieures d'Amazon EMR, EMRFS prend en charge une stratégie de nouvelle tentative alternative basée sur un modèle d'augmentation/diminution multiplicative (AIMD). La stratégie de nouvelle tentative AIMD est particulièrement utile lorsque vous travaillez avec de grands clusters Amazon EMR.

L'AIMD calcule un taux de demandes personnalisé en utilisant les données relatives aux récentes demandes réussies. Cette stratégie réduit le nombre de demandes limitées et le nombre total de tentatives requises par demande.

Pour activer la stratégie de nouvelle tentative AIMD, vous devez définir la propriété `fs.s3.aimd.enabled` sur `true` dans votre configuration `emrfs-site`, comme dans l'exemple suivant.

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "fs.s3.aimd.enabled": "true"
      }
    }
]
```

Pour plus d'informations sur l'utilisation des objets de configuration, consultez Configuration des applications.

Paramètres de nouvelle tentative AIMD avancés

Vous pouvez configurer les propriétés répertoriées dans le tableau suivant pour affiner le comportement des nouvelles tentatives lorsque vous utilisez la stratégie de nouvelle tentative AIMD. Dans la plupart des cas d'utilisation, nous vous recommandons d'utiliser les valeurs par défaut.

Propriétés avancées de la stratégie de nouvelle tentative AIMD

Propriété	Valeur par défaut	Description
`fs.s3.aimd.increaseIncrement`	0.1	Contrôle la rapidité avec laquelle le taux de demandes

Propriété	Valeur par défaut	Description
		augmente lorsque des demandes consécutives aboutissent.
`fs.s3.aimd.reductionFactor`	2	Contrôle la rapidité avec laquelle le taux de demandes diminue lorsqu'Amazon S3 renvoie une réponse 503. Le facteur par défaut de 2 réduit le taux de demandes de moitié.
`fs.s3.aimd.minRate`	0.1	Définit la limite inférieure du taux de demandes lorsque les demandes sont soumises à une limitation soutenue par S3.
`fs.s3.aimd.initialRate`	5500	Définit le taux de demande initial, qui change ensuite en fonction des valeurs que vous spécifiez pour `fs.s3.aimd.increaseIncrement` et `fs.s3.aimd.reductionFactor` . Le taux initial est également utilisé pour les requêtes GET et ajusté proportionnellement (3500/5500) pour les requêtes PUT.
`fs.s3.aimd.adjustWindow`	2	Contrôle la fréquence à laquelle le taux de demandes est ajusté, mesuré en nombre de réponses.

Propriété	Valeur par défaut	Description
`fs.s3.aimd.maxAttempts`	100	Définit le nombre maximum de tentatives pour essayer une demande.

Ajout d'une étape Spark

Vous pouvez utiliser des étapes Amazon EMR pour envoyer des tâches à l'infrastructure Spark installée sur un cluster EMR. Pour plus d'informations, consultez Étapes dans le Guide de gestion d'Amazon EMR. Dans la console et l'interface de ligne de commande, vous faites cela en utilisant une étape d'application Spark, qui exécute le script `spark-submit` comme une étape en votre compte. Avec l'API, vous utilisez une étape pour appeler `spark-submit` à l'aide de `command-runner.jar`.

Pour plus d'informations sur la soumission des applications à Spark, consultez la rubrique Soumission d'applications dans la documentation Apache Spark.

Pour soumettre une étape Spark à l'aide de la console

1. Ouvrez la console Amazon EMR à l'adresse /emr. https://console.aws.amazon.com
2. Dans la Liste de clusters, choisissez le nom de votre cluster.
3. Faites défiler l'affichage jusqu'à la section Étapes et développez-la, puis choisissez Ajouter une étape.
4. Dans la boîte de dialogue Ajouter une étape :
   - Pour Type d'étape, choisissez Spark application.
   - Pour Nom, acceptez le nom par défaut (application Spark) ou saisissez un nouveau nom.
   - Pour le Mode de déploiement, choisissez Client ou Cluster. Le mode client lance le programme pilote sur l'instance principale du cluster, tandis que le mode cluster lance votre programme pilote sur le cluster. Pour le mode client, la sortie du journal du pilote apparaît dans les journaux d'étape, tandis que pour le mode cluster, la sortie du journal du pilote apparaît dans les journaux du premier conteneur YARN. Pour plus d'informations, consultez Présentation du mode cluster dans la documentation Apache Spark.

- Spécifiez le paramètre Spark-submit options (Options Spark-envoi) souhaité. Pour plus d'informations sur les options `spark-submit`, consultez Lancement d'applications avec spark-submit.
- Pour Application location (Emplacement d'application), spécifiez le chemin d'accès local d'URI local ou S3 de l'application.
- Pour Arguments, laissez le champ vide.
- Pour Action sur échec, acceptez l'option par défaut (Continuer).

5. Choisissez Ajouter. L'étape s'affiche dans la console avec le statut En suspens.
6. Le statut de l'étape passe de Pending (En suspens) à Running (En cours d'exécution) puis à Completed (Terminé), au fur et à mesure de son exécution. Pour mettre à jour l'état, choisissez l'icône d'Actualisation au-dessus de la colonne Actions.
7. Les résultats de l'étape se trouvent sur la page Détails de cluster de la console Amazon EMR, à côté de l'étape elle-même et sous Fichiers de journaux si vous avez configuré la journalisation. Vous trouverez le cas échéant les informations sur l'étape dans le compartiment de journaux que vous avez configuré lorsque vous avez lancé le cluster.

Pour soumettre un travail à Spark à l'aide du AWS CLI

Soumettez une étape lorsque vous créez le cluster ou utilisez la sous-commande `aws emr add-steps` dans un cluster existant.

1. Utilisez `create-cluster` comme illustré dans l'exemple suivant.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Add Spark Step Cluster" --release-label emr-7.9.0 --
applications Name=Spark \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=Spark,Name="Spark Program",ActionOnFailure=CONTINUE,Args=[--
class,org.apache.spark.examples.SparkPi,/usr/lib/spark/examples/jars/spark-
examples.jar,10] --use-default-roles
```

Comme alternative, vous pouvez utiliser `command-runner.jar` comme illustré dans l'exemple suivant.

```
aws emr create-cluster --name "Add Spark Step Cluster" --release-label emr-7.9.0 \
--applications Name=Spark --ec2-attributes KeyName=myKey --instance-type m5.xlarge
 --instance-count 3 \
--steps Type=CUSTOM_JAR,Name="Spark Program",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[spark-example,SparkPi,10] --use-default-
roles
```

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

2. Vous pouvez également ajouter des étapes à un cluster déjà en cours d'exécution. Utilisez `add-steps`.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps
 Type=Spark,Name="Spark Program",ActionOnFailure=CONTINUE,Args=[--
class,org.apache.spark.examples.SparkPi,/usr/lib/spark/examples/jars/spark-
examples.jar,10]
```

Comme alternative, vous pouvez utiliser `command-runner.jar` comme illustré dans l'exemple suivant.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps Type=CUSTOM_JAR,Name="Spark
 Program",Jar="command-runner.jar",ActionOnFailure=CONTINUE,Args=[spark-
example,SparkPi,10]
```

Pour soumettre un travail à Spark à l'aide du kit SDK pour Java

1. L'exemple suivant montre comment ajouter une étape à un cluster avec Spark à l'aide de Java.

```
AWSCredentials credentials = new BasicAWSCredentials(accessKey, secretKey);
AmazonElasticMapReduce emr = new AmazonElasticMapReduceClient(credentials);
```

```
StepFactory stepFactory = new StepFactory();
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
AddJobFlowStepsRequest req = new AddJobFlowStepsRequest();
req.withJobFlowId("j-1K48XXXXXXHCB");

List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

HadoopJarStepConfig sparkStepConf = new HadoopJarStepConfig()
   .withJar("command-runner.jar")
   .withArgs("spark-submit","--executor-memory","1g","--
class","org.apache.spark.examples.SparkPi","/usr/lib/spark/examples/jars/spark-
examples.jar","10");

StepConfig sparkStep = new StepConfig()
   .withName("Spark Step")
   .withActionOnFailure("CONTINUE")
   .withHadoopJarStep(sparkStepConf);

stepConfigs.add(sparkStep);
req.withSteps(stepConfigs);
AddJobFlowStepsResult result = emr.addJobFlowSteps(req);
```

2. Affichez les résultats de l'étape en examinant les journaux de l'étape. Vous pouvez le faire dans le AWS Management Console si vous avez activé la journalisation en choisissant Étapes, en sélectionnant votre étape, puis, pour les fichiers journaux, en choisissant `stdout` soit`stderr`. Choisissez Afficher les journaux pour afficher les journaux disponibles.

Remplacement des paramètres de configuration par défaut de Spark

Vous souhaiterez probablement remplacer les valeurs de configuration par défaut de Spark indépendamment pour chaque application. Vous pouvez faire cela lorsque vous soumettez des applications à l'aide d'une étape, ce qui transmet essentiellement des options à `spark-submit`. Par exemple, vous pouvez changer la mémoire allouée à un processus d'exécuteur en modifiant `spark.executor.memory`. Vous pouvez fournir le commutateur `--executor-memory` avec un argument tel que :

```
spark-submit --executor-memory 1g --class org.apache.spark.examples.SparkPi /usr/lib/
spark/examples/jars/spark-examples.jar 10
```

De même, vous pouvez ajuster `--executor-cores` et `--driver-memory`. Dans une étape, vous pouvez fournir les arguments suivants à l'étape :

```
--executor-memory 1g --class org.apache.spark.examples.SparkPi /usr/lib/spark/examples/
jars/spark-examples.jar 10
```

Vous pouvez également ajuster les paramètres qui n'ont peut-être pas de commutateur intégré à l'aide de l'option `--conf`. Pour plus d'informations sur les autres paramètres ajustables, consultez la rubrique Chargement dynamique de propriétés Spark dans la documentation Apache Spark.

Afficher l'historique de l'application Spark

Vous pouvez consulter les détails de Spark, de l'application YARN et de l'interface utilisateur de Tez à l'aide de l'onglet Interfaces utilisateur de l'application de la page détaillée d'un cluster dans la console. Les interfaces utilisateur (UI) des applications Amazon EMR vous permettent de dépanner et d'analyser plus facilement les tâches actives et l'historique des tâches.

Pour plus d'informations, consultez Afficher l'historique de l'application dans le Guide de gestion Amazon EMR.

Accédez au site Web de Spark UIs

Vous pouvez consulter le site Web Spark UIs en suivant les procédures pour créer un tunnel SSH ou créer un proxy dans la section intitulée « Connect to the cluster » du guide de gestion Amazon EMR, puis en accédant au ResourceManager YARN correspondant à votre cluster. Choisissez le lien sous Tracking UI (UI de suivi) pour votre application. Si votre application est en cours d'exécution, vous voyez ApplicationMaster. Vous êtes dirigé vers l'interface utilisateur web du principal de l'application Spark sur le port 20888, quel que soit l'endroit où se trouve le pilote. Le pilote peut être situé sur le nœud primaire du cluster si vous l'exécutez en mode client YARN. Si vous exécutez une application en mode cluster YARN, le pilote se trouve dans le répertoire ApplicationMaster de l'application sur le cluster. Si votre application est terminée, l'historique s'affiche, qui vous permet d'accéder au numéro de port Spark HistoryServer UI 18080 du nœud principal du cluster EMR. Ceci s'applique aux applications déjà terminées. Vous pouvez également accéder à l' HistoryServer interface utilisateur de Spark directement à l'adresse http : *`master-public-dns-name`* //:18080/.

Avec Amazon EMR version 5.25.0 et ultérieure, vous pouvez accéder à l'interface utilisateur du serveur d'historique Spark à partir de la console sans configurer un proxy web via une connexion SSH. Pour plus d'informations, consultez Afficher les interfaces utilisateur d'application persistante.

Utilisation du connecteur Amazon Kinesis Data Streams pour le streaming structuré Spark

Les versions 7.1.0 et supérieures d'Amazon EMR incluent un connecteur Amazon Kinesis Data Streams de streaming structuré Spark dans l'image de publication. Avec ce connecteur, vous pouvez utiliser Spark sur Amazon EMR pour traiter les données stockées dans Amazon Kinesis Data Streams. Le connecteur prend en charge à la fois les types de consommation `GetRecords` (débit partagé) et `SubscribeToShard` (ventilation améliorée). Cette intégration est basée sur le `spark-sql-kinesis-connector`. Pour plus de détails sur la façon de commencer à utiliser le connecteur, consultez le fichier README.

L'exemple suivant montre comment utiliser le connecteur pour lancer une application Spark avec Amazon EMR.

```
spark-submit my_kinesis_streaming_script.py
```

Utilisation de l'intégration d'Amazon Redshift pour Apache Spark avec Amazon EMR

Avec Amazon EMR version 6.4.0 et versions ultérieures, chaque image de version inclut un connecteur entre Apache Spark et Amazon Redshift. Avec ce connecteur, vous pouvez utiliser Spark sur Amazon EMR pour traiter les données stockées dans Amazon Redshift. Pour les versions 6.4.0 à 6.8.0 d'Amazon EMR, l'intégration est basée sur le connecteur open source `spark-redshift`. Pour les versions 6.9.0 et ultérieures d'Amazon EMR, l'intégration Amazon Redshift pour Apache Spark a été migrée de la version communautaire vers une intégration native.

Rubriques

- Lancement d'une application Spark à l'aide de l'intégration Amazon Redshift pour Apache Spark
- Authentification avec l'intégration Amazon Redshift pour Apache Spark
- Lecture et écriture à partir de et vers Amazon Redshift
- Considérations et limites relatives à l'utilisation du connecteur Spark

Lancement d'une application Spark à l'aide de l'intégration Amazon Redshift pour Apache Spark

Pour les versions 6.4 à 6.9 d'Amazon EMR, vous devez utiliser l'option `--jars` ou `--packages` pour spécifier les fichiers JAR suivants que vous souhaitez utiliser. L'option `--jars` spécifie les dépendances stockées localement, dans HDFS ou à l'aide du protocole HTTP/S. Pour connaître les autres emplacements de fichiers pris en charge par l'option `--jars`, consultez la section Gestion avancée des dépendances dans la documentation de Spark. L'option `--packages` spécifie les dépendances stockées dans le référentiel public de Maven.

- `spark-redshift.jar`
- `spark-avro.jar`
- `RedshiftJDBC.jar`
- `minimal-json.jar`

Les versions 6.10.0 et supérieures d'Amazon EMR ne nécessitent pas la dépendance de `minimal-json.jar` et installent automatiquement les autres dépendances sur chaque cluster par défaut. Les exemples suivants montrent comment lancer une application Spark avec l'intégration Amazon Redshift pour Apache Spark.

Amazon EMR 6.10.0 +

L'exemple suivant montre comment lancer une application Spark avec le connecteur `spark-redshift` avec Amazon EMR versions 6.10 et supérieures.

```
spark-submit my_script.py
```

Amazon EMR 6.4.0 - 6.9.x

Pour lancer une application Spark avec le connecteur `spark-redshift` sur les versions 6.4 à 6.9 d'Amazon EMR, vous devez utiliser l'option `--jars` ou `--packages`, comme le montre l'exemple suivant. Notez que les chemins répertoriés avec l'option `--jars` sont les chemins par défaut des fichiers JAR.

```
spark-submit \
  --jars /usr/share/aws/redshift/jdbc/RedshiftJDBC.jar,/usr/share/aws/redshift/
spark-redshift/lib/spark-redshift.jar,/usr/share/aws/redshift/spark-redshift/lib/
spark-avro.jar,/usr/share/aws/redshift/spark-redshift/lib/minimal-json.jar \
```

```
  my_script.py
```

Authentification avec l'intégration Amazon Redshift pour Apache Spark

Utilisation AWS Secrets Manager pour récupérer les informations d'identification et se connecter à Amazon Redshift

L'exemple de code suivant montre comment récupérer les informations d'identification AWS Secrets Manager pour vous connecter à un cluster Amazon Redshift avec l' PySpark interface d'Apache Spark en Python.

```
from pyspark.sql import SQLContext
import boto3

sc = # existing SparkContext
sql_context = SQLContext(sc)

secretsmanager_client = boto3.client('secretsmanager')
secret_manager_response = secretsmanager_client.get_secret_value(
    SecretId='string',
    VersionId='string',
    VersionStage='string'
)
username = # get username from secret_manager_response
password = # get password from secret_manager_response
url = "jdbc:redshift://redshifthost:5439/database?user=" + username + "&password=" +
 password

# Read data from a table
df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("dbtable", "my_table") \
    .option("tempdir", "s3://path/for/temp/data") \
    .load()
```

Utilisation d'IAM pour récupérer les informations d'identification et se connecter à Amazon Redshift

Vous pouvez utiliser le JDBC version 2 fourni par Amazon Redshift pour vous connecter à Amazon Redshift avec le connecteur Spark. Pour utiliser AWS Identity and Access Management (IAM), configurez votre URL JDBC pour utiliser l'authentification IAM. Pour vous connecter à un cluster Redshift depuis Amazon EMR, vous devez donner à votre rôle IAM l'autorisation de récupérer des informations d'identification IAM temporaires. Attribuez les autorisations suivantes à votre rôle IAM pour qu'il puisse récupérer les informations d'identification et exécuter les opérations Amazon S3.

- Redshift : GetClusterCredentials (pour les clusters Amazon Redshift provisionnés)
- Redshift : DescribeClusters (pour les clusters Amazon Redshift provisionnés)
- Redshift : GetWorkgroup (pour les groupes de travail Amazon Redshift Serverless)
- Redshift : GetCredentials (pour les groupes de travail Amazon Redshift Serverless)
- s3 : GetBucket
- s3 : GetBucketLocation
- s3 : GetObject
- s3 : PutObject
- s3 : GetBucketLifecycleConfiguration

Pour plus d'informations sur `GetClusterCredentials`, consultez Stratégies de ressources pour `GetClusterCredentials`.

Vous devez également vous assurer qu'Amazon Redshift peut assumer le rôle IAM pendant les opérations `COPY` et `UNLOAD`.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "redshift.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
```

```
}
```

L'exemple suivant utilise l'authentification IAM entre Spark et Amazon Redshift :

```
from pyspark.sql import SQLContext
import boto3

sc = # existing SparkContext
sql_context = SQLContext(sc)

url = "jdbc:redshift:iam://redshift-host:redshift-port/db-name"
iam_role_arn = "arn:aws:iam::account-id:role/role-name"

# Read data from a table
df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("aws_iam_role", iam_role_arn) \
    .option("dbtable", "my_table") \
    .option("tempdir", "s3a://path/for/temp/data") \
    .mode("error") \
    .load()
```

Lecture et écriture à partir de et vers Amazon Redshift

Les exemples de code suivants permettent de lire et PySpark d'écrire des exemples de données depuis et vers une base de données Amazon Redshift à l'aide de l'API de source de données et de SparkSQL.

Data source API

PySpark À utiliser pour lire et écrire des exemples de données depuis et vers une base de données Amazon Redshift avec une API de source de données.

```
import boto3
from pyspark.sql import SQLContext

sc = # existing SparkContext
sql_context = SQLContext(sc)

url = "jdbc:redshift:iam://redshifthost:5439/database"
```

```
aws_iam_role_arn = "arn:aws:iam::accountID:role/roleName"

df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("dbtable", "tableName") \
    .option("tempdir", "s3://path/for/temp/data") \
    .option("aws_iam_role", "aws_iam_role_arn") \
    .load()

df.write \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("dbtable", "tableName_copy") \
    .option("tempdir", "s3://path/for/temp/data") \
    .option("aws_iam_role", "aws_iam_role_arn") \
    .mode("error") \
    .save()
```

SparkSQL

PySpark À utiliser pour lire et écrire des exemples de données depuis et vers une base de données Amazon Redshift avec SparkSQL.

```
import boto3
import json
import sys
import os
from pyspark.sql import SparkSession

spark = SparkSession \
    .builder \
    .enableHiveSupport() \
    .getOrCreate()

url = "jdbc:redshift:iam://redshifthost:5439/database"
aws_iam_role_arn = "arn:aws:iam::accountID:role/roleName"

bucket = "s3://path/for/temp/data"
tableName = "tableName" # Redshift table name

s = f"""CREATE TABLE IF NOT EXISTS {tableName} (country string, data string)
    USING io.github.spark_redshift_community.spark.redshift
```

```
    OPTIONS (dbtable '{tableName}', tempdir '{bucket}', url '{url}', aws_iam_role
 '{aws_iam_role_arn}' ); """

spark.sql(s)

columns = ["country" ,"data"]
data = [("test-country","test-data")]
df = spark.sparkContext.parallelize(data).toDF(columns)

# Insert data into table
df.write.insertInto(tableName, overwrite=False)
df = spark.sql(f"SELECT * FROM {tableName}")
df.show()
```

Considérations et limites relatives à l'utilisation du connecteur Spark

- Nous vous recommandons d'activer le protocole SSL pour la connexion JDBC entre Spark sur Amazon EMR et Amazon Redshift.
- À titre de bonne pratique, nous vous recommandons de gérer les informations d'identification du cluster Amazon Redshift dans AWS Secrets Manager . Voir Utiliser AWS Secrets Manager pour récupérer les informations d'identification pour se connecter à Amazon Redshift pour un exemple.
- Nous vous recommandons de transmettre un rôle IAM à l'aide du paramètre `aws_iam_role` pour le paramètre d'authentification Amazon Redshift.
- L'URI `tempdir` renvoie à un emplacement Amazon S3. Ce répertoire temporaire n'est pas nettoyé automatiquement et peut donc entraîner des coûts supplémentaires.
- Tenez compte des recommandations suivantes pour Amazon Redshift :
  - Nous vous recommandons de bloquer l'accès public au cluster Amazon Redshift.
  - Nous vous recommandons d'activer la journalisation des audits d'Amazon Redshift.
  - Nous vous recommandons d'activer le chiffrement au repos d'Amazon Redshift.
- Tenez compte des recommandations suivantes pour Amazon S3 :
  - Nous vous recommandons de bloquer l'accès public aux compartiments Amazon S3.
  - Nous vous recommandons d'utiliser le chiffrement côté serveur Amazon S3 pour chiffrer les compartiments Amazon S3 utilisés.
  - Nous vous recommandons d'utiliser les politiques de cycle de vie d'Amazon S3 pour définir les règles de conservation du compartiment Amazon S3.

- Amazon EMR vérifie toujours le code importé à partir d'une source ouverte dans l'image. Pour des raisons de sécurité, nous ne prenons pas en charge les méthodes d'authentification suivantes de Spark vers Amazon S3 :
  - Configuration des clés AWS d'accès dans la classification `hadoop-env` de configuration
  - Encodage des clés AWS d'accès dans l'`tempdir`URI

Pour plus d'informations sur l'utilisation du connecteur et les paramètres qu'il prend en charge, consultez les ressources suivantes :

- Intégration d'Amazon Redshift pour Apache Spark dans le Guide de gestion Amazon Redshift
- Le référentiel communautaire `spark-redshift` sur Github

Historique des versions de Spark

Le tableau suivant répertorie la version de Spark incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la rubrique Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

> **Important**
>
> Apache Spark version 2.3.1, disponible à partir de la version 5.16.0 d'Amazon EMR, corrige CVE-2018-8024 et CVE-2018-1334. Nous vous recommandons de migrer les versions antérieures de Spark vers la version 2.3.1 ou ultérieure.

Informations sur la version de Spark

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-7.9.0	3.5.5	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
		kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7,8.0	3.5.4	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-7,7.0	3.5.3	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.6.0	3.5.3	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-7.5.0	3.5.2	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.4.0	3.5.2	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-7.3.0	3.5.1	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.2.0	3.5.1	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.36.2	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.1.0	3.5.0	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-7.0.0	3.5.0	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.15.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.14.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.13.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.12.0	3.4.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.11.1	3.3.2	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.11.0	3.3.2	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.10.1	3.3.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.10.0	3.3.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.9.1	3.3.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.9.0	3.3.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.8.1	3.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.8.0	3.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.7.0	3.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.36.1	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.36.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.6.0	3.2.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.35.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.5.0	3.1.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.4.0	3.1.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.3.1	3.1.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.3.0	3.1.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.2.1	3.0.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.2.0	3.0.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.1.1	3.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.1.0	3.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-6.0.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.0.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.34.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.33.1	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.33.0	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.32.1	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.32.0	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.31.1	2.4.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.31.0	2.4.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.30.2	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.30.1	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.30.0	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-notebook-env, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
mr-5.29.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.28.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.28.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.27.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.27.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.26.0	2.4.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.25.0	2.4.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.24.1	2.4.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.24.0	2.4.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.23.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.23.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.22.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.21.2	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.21.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.21.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.20.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.20.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.19.1	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.19.0	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.18.1	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.18.0	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.17.2	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.17.1	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.17.0	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.16.1	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.16.0	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.15.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.15.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.14.2	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.14.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.14.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.13.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.13.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.12.3	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.12.2	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.12.1	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.12.0	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.4	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.11.3	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.2	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.11.1	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.0	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.10.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.10.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.9.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.9.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.8.3	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.8.2	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.8.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.8.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.7.1	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.7.0	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.6.1	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.6.0	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.5.4	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.3	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.2	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.5.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.4.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.4.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.3.2	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.3.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.3.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.3	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.2	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.2.1	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.0	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.1.1	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.1.0	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.3	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.2	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-5.0.1	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.0	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.6	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-4.9.5	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.4	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.3	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-4.9.2	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.1	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.5	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-4.8.4	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.3	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.2	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-4.8.1	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.0	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.4	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-4.7.3	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.2	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.1	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-4.7.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.6.1	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.6.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-4.5.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.4.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.3.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Étiquette de version Amazon EMR	Version de Spark	Composants installés avec Spark
emr-4.2.0	1.5.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.1.0	1.5.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.0.0	1.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Apache Sqoop

> **Note**
>
> La dernière version à inclure Apache Sqoop était la version 7.4 d'Amazon EMR. Apache Sqoop sera supprimé des versions 7.5 et supérieures d'Amazon EMR.

Apache Sqoop est un outil de transfert de données entre Amazon S3, Hadoop, HDFS et les bases de données SGBDR. Pour plus d'informations, consultez le site web Apache Sqoop. Sqoop est inclus dans les versions Amazon EMR 5.0.0 et supérieures. Les versions antérieures incluent Sqoop en tant qu'application sandbox. Pour de plus amples informations, veuillez consulter Versions Amazon EMR 4.x.

Rubriques

- Informations sur la version de Sqoop
- Éléments à prendre en compte avec Sqoop sur Amazon EMR
- Historique des versions de Sqoop

Informations sur la version de Sqoop

Version Sqoop pour la version 7.4.

Le tableau suivant répertorie la version de Sqoop incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants qu'Amazon EMR installe avec Sqoop.

Informations sur la version de Spark pour emr-7.4.0

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-7.4.0	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Version Sqoop pour 6.15.0

Le tableau suivant répertorie la version de Sqoop incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Sqoop.

Pour connaître la version des composants installés avec Sqoop dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version de Sqoop pour emr-6.15.0

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-6.15.0	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Version Sqoop pour 5.36.2

Le tableau suivant répertorie la version de Sqoop incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Sqoop.

Pour la version des composants installés avec Sqoop dans cette version, voir la version 5.36.2 Versions des composants.

Informations sur la version de Sqoop pour emr-5.36.2

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.36.2	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Éléments à prendre en compte avec Sqoop sur Amazon EMR

Tenez compte des éléments suivants lorsque vous exécutez Sqoop sur Amazon EMR.

Utilisation de Sqoop avec intégration HCatalog

Sqoop sur Amazon EMR prend en charge l'intégration de Sqoop. HCatalog Lorsque vous utilisez Sqoop pour écrire une sortie dans une HCatalog table dans Amazon S3, désactivez l'écriture directe Amazon EMR en définissant les `mapred.output.direct.NativeS3FileSystem` propriétés et sur. `mapred.output.direct.EmrFileSystem` `false` Pour de plus amples informations, veuillez consulter En utilisant HCatalog. Vous pouvez utiliser les commandes Hadoop `-D mapred.output.direct.NativeS3FileSystem=false` et `-D mapred.output.direct.EmrFileSystem=false`. Si vous ne désactivez pas l'écriture directe, aucune erreur ne se produit, mais la table est créée dans Amazon S3 et aucune donnée n'est écrite.

Prise en charge des bases de données et JDBC par Sqoop

Par défaut, Sqoop dispose d'un pilote MariaDB et PostgreSQL installé. Le pilote PostgreSQL installé pour Sqoop fonctionne uniquement pour PostgreSQL 8.4. Pour installer un autre jeu de connecteurs JDBC pour Sqoop, connectez-vous au nœud principal de cluster et installez-les dans `/usr/lib/sqoop/lib`. Les liens suivants se rapportent à divers connecteurs JDBC :

- MariaDB : A propos de MariaDB Connector/J.
- PostgreSQL : Pilote JDBC PostgreSQL.
- SQLServer: Téléchargez le pilote Microsoft JDBC pour SQL Server.
- MySQL : Télécharger le connecteur/J
- Oracle : Obtenir les pilotes et UCP JBDC à partir du référentiel Oracle Maven

Les bases de données prises en charge pour Sqoop sont répertoriées à l'adresse suivante, http://sqoop.apache.org/docs/*version*/SqoopUserGuide.html #_supported_databases, où se *version* trouve la version de Sqoop que vous utilisez, par exemple 1.4.6. Si la chaîne de connexion JDBC ne correspond pas à celles figurant dans cette liste, vous devez spécifier un pilote.

Par exemple, vous pouvez exporter vers une table de base de données Amazon Redshift avec la commande suivante (pour JDBC 4.1) :

```
sqoop export --connect jdbc:redshift://$MYREDSHIFTHOST:5439/mydb --
table mysqoopexport --export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver
 com.amazon.redshift.jdbc41.Driver --username master --password Mymasterpass1
```

Vous pouvez utiliser les chaînes de connexion MariaDB et MySQL, mais si vous spécifiez la chaîne de connexion MariaDB, vous devez spécifier le pilote :

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table mysqoopexport --
export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver org.mariadb.jdbc.Driver --
username master --password Mymasterpass1
```

Si vous utilisez un chiffrement SSL (Secure Socket Layer) pour accéder à votre base de données, vous devez utiliser un URI JDBC comme dans l'exemple d'exportation Sqoop suivant :

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb?
verifyServerCertificate=false&useSSL=true&requireSSL=true --table mysqoopexport --
```

```
export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver org.mariadb.jdbc.Driver --
username master --password Mymasterpass1
```

Pour plus d'informations sur le chiffrement SSL dans RDS, consultez la section Utilisation de SSL pour chiffrer une connexion à une instance de base de données dans le Guide de l'utilisateur Amazon RDS.

Pour plus d'informations, consultez la documentation Apache Sqoop.

Sécurisation de votre mot de passe

Différentes méthodes permettent de sécuriser votre mot de passe :

Java KeyStore

La méthode préférée consiste à chiffrer le mot de passe à l'aide d'un code Java KeyStore (JKS), éliminant ainsi le besoin de le stocker dans un format lisible.

1. Créez un alias de mot de passe. Lorsque vous y êtes invité, entrez le mot de passe que vous utilisez pour accéder à la base de données.

   ```
   hadoop credential create mydb.password.alias -provider jceks://hdfs/user/root/
   mysql.password.jceks
   ```

2. Utilisez l'alias de mot de passe pour lancer la tâche Sqoop :

   ```
   sqoop export -Dhadoop.security.credential.provider.path=jceks://hdfs/user/
   root/mysql.password.jceks --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table
    mysqoopexport --export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver
    org.mariadb.jdbc.Driver --username master --password-alias mydb.password.alias
   ```

--password-file

Vous pouvez utiliser la commande `--password-file` pour transmettre le mot de passe dans un fichier comme illustré dans l'exemple suivant :

1. Créez un fichier contenant le mot de passe :

   ```
   echo -n 'Mymasterpass1' > /home/hadoop/mysql-pass.password
   ```

2. Utilisez le fichier pour lancer la tâche Sqoop :

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table
 mysqoopexport --export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver
 org.mariadb.jdbc.Driver --username master --password-file /home/hadoop/mysql-
pass.password
```

-P

Vous pouvez utiliser la commande `-P` pour entrer le mot de passe dans une invite comme illustré dans l'exemple suivant :

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table mysqoopexport --
export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver org.mariadb.jdbc.Driver
 --username master -P
```

Historique des versions de Sqoop

Le tableau suivant répertorie la version de Sqoop incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version de Sqoop

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-7.4.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
		timeline-server, mariadb-server, sqoop-client
emr-7.3.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-7.2.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.36.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-7.1.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-7.0.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.15.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-6.14.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.13.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-6.12.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.11.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-6.11.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.10.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-6.10.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.9.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-6.9.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.8.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-6.8.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.7.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.36.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.36.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-6.6.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.35.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-6.5.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.4.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-6.3.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.3.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-6.2.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.2.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-6.1.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.1.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.34.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.33.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.33.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.32.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.32.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.31.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.31.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.30.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.30.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.30.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
mr-5.29.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.28.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.28.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.27.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.27.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.26.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.25.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.24.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.24.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.23.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.23.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.22.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.21.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.21.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.21.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.20.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.20.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.19.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.19.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.18.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.18.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.17.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.17.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.17.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.16.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.16.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.15.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.15.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.14.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.14.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.14.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.13.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.13.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.12.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.12.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.12.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.12.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.11.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.11.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.10.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.10.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.9.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.9.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.8.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.8.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.8.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.8.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.7.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.7.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.6.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.6.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.5.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.5.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.5.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.5.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.5.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.4.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.4.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.3.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.3.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.3.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.2.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.2.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.2.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.2.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.1.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.1.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.0.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.0.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.0.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Étiquette de version Amazon EMR	Version Sqoop	Composants installés avec Sqoop
emr-5.0.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

TensorFlow

TensorFlow est une bibliothèque de calcul symbolique open source pour l'intelligence de machine et les applications de deep learning. Pour plus d'informations, consultez le TensorFlow site web. TensorFlow est disponible avec Amazon EMR version 5.17.0 ou ultérieure.

Le tableau suivant répertorie la version TensorFlow incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants avec Amazon EMR installe. TensorFlow

Pour connaître la version des composants installés avec TensorFlow cette version, consultez les versions des composants de la version 7.9.0.

TensorFlow informations de version pour emr-7.9.0

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-7.9.0	TensorFlow 2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Le tableau suivant répertorie la version TensorFlow incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants avec Amazon EMR installe. TensorFlow

Pour connaître la version des composants installés avec TensorFlow cette version, consultez les versions des composants de la version 6.15.0.

TensorFlow informations sur la version pour emr-6.15.0

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-6.15.0	TensorFlow 2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Le tableau suivant répertorie la version TensorFlow incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants avec Amazon EMR installe. TensorFlow

Pour connaître la version des composants installés avec TensorFlow cette version, consultez les versions des composants de la version 5.36.2.

TensorFlow informations de version pour emr-5.36.2

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.36.2	TensorFlow 2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

TensorFlow compilations par type d' EC2 instance Amazon

Amazon EMR utilise différents builds de la TensorFlow bibliothèque en fonction du type d'instance que vous choisissez pour votre cluster. Amazon EMR prend en charge les clusters dotés de types d'instances aarch64 (Graviton) TensorFlow pour EMR-7.5.0 et versions ultérieures. Le tableau suivant répertorie les builds par type d'instance.

EC2 types d'instances	TensorFlow construire
M5 et C5	Tensorflow 2.16.1 avec optimisation Intel MKL
P2, P4D, P5, G4DN, G5, G6 et GR6	Tensorflow 2.16.1 avec CUDA 12.3, cuDNN 8.9.7.29
P3, P3DN, G3 et G3S	Tensorflow 2.16.1 avec CUDA 12.3, cuDNN 8.9.7.29, NCCL 2.20.3-1 Nvidia NCCL est disponible uniquement sur les instances P3. Contrat de Licence Utilisateur Final (CLUF) : en utilisant les composants Nvidia sur Amazon EMR, vous acceptez les termes et conditions décrites dans le CLUF de produit.
Tous les autres sauf les instances de Graviton	Tensorflow 2.16.1

Sécurité

En plus de conseils dans Utilisation TensorFlow sécurisée, nous vous recommandons de lancer votre cluster dans un sous-réseau privé pour vous aider à limiter l'accès aux sources fiables. Pour plus d'informations, consultez Options d'Amazon VPC dans le Guide de gestion Amazon EMR.

En utilisant TensorBoard

TensorBoard est une suite d'outils de visualisation pour les TensorFlow programmes. Pour plus d'informations, consultez TensorBoard: Apprentissage visualisé sur le site web Tensorflow.

Pour l'utiliser TensorBoard avec Amazon EMR, vous devez commencer TensorBoard sur le nœud principal du cluster.

Pour utiliser tensorboard avec Tensorflow sur Amazon EMR

1. Connectez-vous au nœud maître du cluster à l'aide de SSH. Pour plus d'informations, consultez Connexion au nœud principal à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
2. Tapez la commande suivante pour démarrer Tensorboard sur le nœud principal. Remplacez */my/log/directory* par un répertoire sur le nœud principal où vous avez généré et résumé des données stockées à l'aide d'un enregistreur de résumé.

   Amazon EMR 5.19.0 and later

   ```
   python3 -m tensorboard.main --logdir=/home/hadoop/tensor --bind_all
   ```

   Amazon EMR 5.18.1 and earlier

   ```
   python3 -m tensorboard.main --logdir=/my/log/dir
   ```

   Par défaut, le nœud principal héberge TensorBoard en utilisant le port 6006 et le nom DNS public principal. Une fois que vous avez démarré TensorBoard, la sortie de la ligne de commande présente l'URL à laquelle vous pouvez vous connecter TensorBoard, comme illustré dans l'exemple suivant :

   ```
   TensorBoard 2.16.1 at http://master-public-dns-name:6006 (Press CTRL+C to quit)
   ```
3. Configurez l'accès aux interfaces web sur le nœud principal à partir de clients de confiance. Pour plus d'informations, consultez Affichage des interfaces web hébergées sur les clusters Amazon EMR dans le Guide de gestion Amazon EMR.
4. Ouvert TensorBoard à`http://master-public-dns-name:6006`.

TensorFlow historique des versions

Le tableau suivant répertorie la version TensorFlow incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

TensorFlow informations sur la version

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-7.9.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.8.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7,7.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.6.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
		hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.5.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.4.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-7.3.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.2.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.36.2	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-7.1.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.0.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.15.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-6.14.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.13.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.12.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-6.11.1	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.11.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.10.1	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-6.10.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.9.1	2.10.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.9.0	2.10.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-6.8.1	2.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.8.0	2.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.7.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.36.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.36.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.6.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.35.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.5.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.4.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-6.3.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.3.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.2.1	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-6.2.0	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.1.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.1.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-6.0.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.0.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.34.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.33.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.33.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.32.1	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.32.0	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.31.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.31.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.30.2	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.30.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.30.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
mr-5.29.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.28.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.28.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.27.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.27.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.26.0	1.13.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.25.0	1.13.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.24.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.24.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.23.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.23.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.22.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.21.2	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.21.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.21.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.20.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.20.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.19.1	1.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.19.0	1.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.18.1	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.18.0	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Étiquette de version Amazon EMR	TensorFlow Version	Composants installés avec TensorFlow
emr-5.17.2	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.17.1	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.17.0	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Apache Tez

Apache Tez est un cadre qui crée un graphe orienté acyclique (DAG) complexe de tâches pour le traitement des données. Vous pouvez l'utiliser comme alternative à Hadoop dans certains cas MapReduce d'utilisation. Par exemple, vous pouvez exécuter des flux de travail Pig and Hive avec Hadoop MapReduce ou vous pouvez utiliser Tez comme moteur d'exécution. Pour plus d'informations, consultez https://tez.apache.org/. Les versions Amazon EMR 4.7.0 et ultérieures incluent Tez.

Le tableau suivant répertorie la version de Tez incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants qu'Amazon EMR installe avec Tez.

Pour la version des composants installés avec Tez dans cette version, voir la version 7.9.0 Versions des composants.

Informations sur la version de Tez pour emr-7.9.0

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-7.9.0	Tez 0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Le tableau suivant répertorie la version de Tez incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants qu'Amazon EMR installe avec Tez.

Pour connaître la version des composants installés avec Tez dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version de Tez pour emr-6.15.0

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-6.15.0	Tez 0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Le tableau suivant répertorie la version de Tez incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants qu'Amazon EMR installe avec Tez.

Pour la version des composants installés avec Tez dans cette version, voir la version 5.36.2 Versions des composants.

Informations sur la version de Tez pour emr-5.36.2

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.36.2	Tez 0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Rubriques

- Création d'un cluster avec Tez
- Configuration de Tez
- Interface utilisateur web Tez
- Serveur de chronologie
- Historique des versions de Tez

Création d'un cluster avec Tez

Pour installer Tez, choisissez Apache Tez comme application lorsque vous créez votre cluster.

Pour créer un cluster avec Tez installé à l'aide de la console

1. Accédez à la nouvelle console Amazon EMR et sélectionnez Changer pour l'ancienne console depuis le menu latéral. Pour plus d'informations sur ce qu'implique le passage à l'ancienne console, consultez la rubrique Utilisation de l'ancienne console.
2. Choisissez Créer un cluster et Go to advanced options (Aller aux options avancées).
3. Dans Configuration des logiciels, choisissez une version de emr-4.7.0 ou supérieure.
4. Sélectionnez Tez avec d'autres applications que vous voulez faire installer par Amazon EMR.
5. Sélectionnez d'autres options si nécessaire, puis choisissez Create cluster (Créer le cluster).

Pour créer un cluster avec Tez à l'aide du AWS CLI

- Utilisez la commande `create-cluster` avec l'option `-- applications` pour spécifier Tez. L'exemple suivant crée un cluster avec Tez installé.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --name "Cluster with Tez" --release-label emr-7.9.0 \
--applications Name=Tez --ec2-attributes KeyName=myKey \
```

```
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

Configuration de Tez

Vous pouvez personnaliser Tez en définissant des valeurs à l'aide de la classification de configuration `tez-site`, qui configure les paramètres dans le fichier de configuration `tez-site.xml`. Pour plus d'informations, consultez TezConfigurationla documentation d'Apache Tez. Pour modifier Hive ou Pig afin d'utiliser le moteur d'exécution Tez, utilisez les classifications de configuration `hive-site` et `pig-properties` correspondantes. Des exemples sont présentés ci-après.

Exemple de configuration

Example Exemple : Personnalisation du niveau d'enregistrement racine de Tez et Configuration de Tez comme moteur d'exécution pour Hive et Pig

L'exemple de commande `create-cluster` ci-dessous crée un cluster avec Tez, Hive et Pig installés. La commande fait référence à un fichier stocké dans Amazon S3, `myConfig.json`, qui spécifie les propriétés pour la classification `tez-site` qui définit `tez.am.log.level` sur `DEBUG` et définit le moteur d'exécution Tez pour Hive et Pig en utilisant les classifications de configuration `hive-site` et `pig-properties`.

> **Note**
>
> Les caractères de continuation de ligne Linux (\) sont inclus pour des raisons de lisibilité. Ils peuvent être supprimés ou utilisés dans les commandes Linux. Pour Windows, supprimez-les ou remplacez-les par un caret (^).

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Tez Name=Hive Name=Pig --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 \
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json --
use-default-roles
```

Vous trouverez ci-dessous des exemples de contenus `myConfig.json`.

```
[
```

```
    {
      "Classification": "tez-site",
      "Properties": {
        "tez.am.log.level": "DEBUG"
      }
    },
    {
      "Classification": "hive-site",
      "Properties": {
        "hive.execution.engine": "tez"
      }
    },
    {
      "Classification": "pig-properties",
      "Properties": {
        "exectype": "tez"
      }
    }
  ]
```

> **Note**
>
> Avec la version 5.21.0 et ultérieures d'Amazon EMR, vous permet de remplacer les configurations de cluster et de spécifier des classifications de configuration supplémentaires pour chaque groupe d'instances dans un cluster en cours d'exécution. Pour ce faire, utilisez la console Amazon EMR, le AWS Command Line Interface (AWS CLI) ou le AWS SDK. Pour plus d'informations, consultez Fourniture d'une configuration pour un groupe d'instances dans un cluster en cours d'exécution.

Ouverture fractionnée asynchrone Tez

Lorsque le chemin de la table contient un grand nombre de petits fichiers et qu'une requête tente de tous les lire, chaque petit fichier correspondant à chaque division est combiné en une unité de fractionnement groupée Tez. Un seul mappeur traite ensuite la division groupée unique Tez. Puisque l'exécution est synchrone, chaque division de la division groupée est traitée une par une. Les objets `RecordReader` doivent donc traiter les divisions de manière synchrone.

Amazon EMR 6.15.0 introduit des configurations que vous pouvez spécifier pour ouvrir de manière asynchrone les divisions d'entrée dans une division groupée Tez. La fonctionnalité a été initiée par

TEZ-4397, mais a connu des régressions dans OSS Hive. EMR Hive a corrigé les régressions et les bogues supplémentaires dans le tableau Hive ACID. Cette amélioration permet d'accélérer les performances des requêtes de lecture lorsqu'il existe un grand nombre de divisions d'entrée dans une seule division groupée Tez.

Nom	Classification	Description
`tez.grouping.split .init.threads`	`tez-site`	Spécifie le nombre de threads de démon que Tez utilise pour pré-initier les `RecordReaders` et les divisions ouvertes. Pour les tables ACID, la valeur maximale prise en charge pour `tez.grouping.split .init.threads` est 1.
`tez.grouping.split .init.recordreaders`	`tez-site`	Spécifie le nombre de `RecordReaders` dont la pré-initialisation doit être conservée par les threads de démon. Cela peut être utile lorsque la division groupée Tez contient un grand nombre de `InputSplits` . L'initial isation de `RecordReaders` pour traiter ces divisions d'entrée peut être effectuée de manière asynchrone avec des threads de démon au lieu d'un traitement séquentiel.

Remarques relatives à la configuration :

Objet	Détails
Paramètres de configuration recommandés	Il est recommandé de définir les paramètres de configuration ci-dessus sur les valeurs que vous souhaitez dans `hive-site` les deux`tez-site`.
Valeurs correspondantes	Les valeurs des paramètres de configuration doivent être identiques dans les deux versions `hive-site` et`tez-site`.
Recommandation LLAP	Il n'est pas recommandé d'utiliser cette fonctionnalité lorsque LLAP est activé.

Analyse comparative de l’ouverture fractionnée asynchrone Tez

Nous avons utilisé les environnements et configurations suivants pour procéder à l’analyse comparative de la capacité d’ouverture fractionnée asynchrone Tez :

- Environnement d’évaluation – Cluster Amazon EMR avec 1 nœud primaire utilisant m5.16xlarge et 16 nœuds principaux utilisant m5.16xlarge.
- Configurations d’évaluation – Pour simuler le scénario d’analyse comparative dans lequel un grand nombre de divisions d’entrée se trouvent dans une seule division groupée Tez, la valeur `tez.grouping.split-count` est définie sur `1`.
- Table utilisée pour l’analyse comparative – La table contient 200 partitions, chaque partition contenant un seul fichier. L’analyse comparative est effectuée lorsque cette table contient des fichiers CSV et lorsque cette table contient des fichiers parquet. Requête Hive pour l’analyse comparative : `SELECT COUNT(*)` à dix reprises dans la table, et prenez le temps d’exécution moyen.
- Configurations pour activer l’ouverture fractionnée asynchrone Tez :
  - `tez.grouping.split.init.threads` = `4`

- `tez.grouping.split.init.recordreaders` = `10`

Jeux de données	Fonctionnalité désactivée (référence)	Fonctionnalité activée	Amélioration
Jeu de données CSV	90,26 secondes	79,20 secondes	12,25 %
Jeu de données Parquet	54,67 secondes	42,23 secondes	22,75 %

Interface utilisateur web Tez

Tez a sa propre interface utilisateur web. Pour afficher l'interface utilisateur Web, consultez l'URL suivante.

```
http://masterDNS:8080/tez-ui
```

Pour activer l'onglet Hive Queries sur l'interface utilisateur Web de Tez, définissez la configuration suivante.

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.exec.pre.hooks": "org.apache.hadoop.hive.ql.hooks.ATSHook",
      "hive.exec.post.hooks": "org.apache.hadoop.hive.ql.hooks.ATSHook",
      "hive.exec.failure.hooks": "org.apache.hadoop.hive.ql.hooks.ATSHook"
    }
  }
]
```

Vous pouvez également consulter les détails de l'interface utilisateur des applications Tez, Spark et YARN à l'aide des liens figurant dans l'onglet Interfaces utilisateur de l'application de la page détaillée d'un cluster dans la console. Les interfaces utilisateur (UI) des applications Amazon EMR sont hébergées hors cluster et sont disponibles après la fin du cluster. Elles ne nécessitent pas que vous établissiez une connexion SSH ou un proxy Web, ce qui vous permet de résoudre plus facilement les problèmes et d'analyser les tâches actives et l'historique des tâches.

Pour plus d'informations, consultez Afficher l'historique de l'application dans le Guide de gestion Amazon EMR.

Serveur de chronologie

Le Serveur de chronologie YARN est configuré pour s'exécuter lorsque Tez est installé. Pour afficher les tâches soumises via Tez ou des moteurs MapReduce d'exécution à l'aide du serveur Timeline, consultez l'interface utilisateur Web à l'aide de l'URL`http://master-public-DNS:8188`. Pour plus d'informations, consultez Affichage des interfaces web hébergées sur les clusters Amazon EMR dans le Guide de gestion Amazon EMR.

Historique des versions de Tez

Le tableau suivant répertorie la version de Tez incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version de Tez

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-7.9.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7,8.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
		namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7,7.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.6.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-7.5.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.4.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.3.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-7.2.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-5.36.2	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-7.1.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-7.0.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.15.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.14.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-6.13.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.12.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.11.1	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-6.11.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.10.1	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.10.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-6.9.1	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.9.0	0,1,2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.8.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-6.8.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.7.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.36.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.36.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.6.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.35.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-6.5.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.4.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.3.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-6.3.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.2.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.2.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-6.1.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.1.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.0.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-6.0.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.34.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.33.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.33.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.32.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.32.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.31.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.31.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.30.2	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.30.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.30.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
mr-5.29.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.28.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.28.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.27.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.27.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.26.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.25.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.24.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.24.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.23.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.23.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.22.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.21.2	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.21.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.21.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.20.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.20.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.19.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.19.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.18.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.18.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.17.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.17.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.17.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.16.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.16.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.15.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.15.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.14.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.14.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.14.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.13.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.13.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.12.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.11.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.11.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.10.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.10.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.9.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.9.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.8.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.8.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.8.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.8.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.7.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.7.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.6.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.6.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.5.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.5.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.4.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.4.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.3.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.3.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.3.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.2.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.2.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.2.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.2.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.1.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.1.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.0.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.0.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.0.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-5.0.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.6	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.5	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-4.9.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-4.9.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.5	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-4.8.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-4.8.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.4	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.3	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Étiquette de version Amazon EMR	Version Tez	Composants installés avec Tez
emr-4.7.2	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.1	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.0	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Notes de publication de Tez par version

Rubriques

- Amazon EMR 7.8.0 - Notes de mise à jour de Tez
- Amazon EMR 7.7.0 - Notes de mise à jour de Tez
- Amazon EMR 7.6.0 - Notes de mise à jour de Tez
- Amazon EMR 7.5.0 - Notes de mise à jour de Tez
- Amazon EMR 7.4.0 - Notes de mise à jour de Tez
- Amazon EMR 7.2.0 - Notes de mise à jour de Tez
- Amazon EMR 7.1.0 - Notes de mise à jour de Tez
- Amazon EMR 7.0.0 - Notes de mise à jour de Tez
- Amazon EMR 6.15.0 – Notes de mise à jour de Tez
- Amazon EMR 6.14.0 - Notes de mise à jour de Tez
- Amazon EMR 6.13.0 - Notes de mise à jour de Tez
- Amazon EMR 6.12.0 - Notes de mise à jour de Tez
- Amazon EMR 6.11.0 - Notes de mise à jour de Tez
- Amazon EMR 6.10.0 - Notes de mise à jour de Tez
- Amazon EMR 6.9.0 - Notes de mise à jour de Tez
- Amazon EMR 6.8.0 - Notes de mise à jour de Tez
- Amazon EMR 6.7.0 - Notes de mise à jour de Tez
- Amazon EMR 6.6.0 - Notes de mise à jour de Tez

Amazon EMR 7.8.0 - Notes de mise à jour de Tez

Amazon EMR 7.8.0 - Testez les modifications

Type	Description
Amélioration	Version 1.78 améliorée de Bouncycastle pour correspondre à Hadoop

Amazon EMR 7.7.0 - Notes de mise à jour de Tez

Amazon EMR 7.7.0 - Testez les modifications

Type	Description
Amélioration	La `tez.task.relaxed.locality` propriété d'Apache Tez contrôle si la planification des tâches respecte strictement les contraintes de localité des données (localité du rack et du nœud). Lorsqu'il est défini sur `true` (par défaut dans EMR-7.6 et versions ultérieures), Tez n'applique pas la localité, ce qui permet d'attribuer des tâches à n'importe quel conteneur disponible, ce qui améliore l'utilisation des ressources et réduit les temps d'attente dans les clusters très fréquentés.

Amazon EMR 7.6.0 - Notes de mise à jour de Tez

Amazon EMR 7.6.0 - Testez les modifications

Type	Description
Amélioration	Les contraintes de localisation des racks et des nœuds ne seront pas prises en compte lors de la demande de conteneur pour une tâche en remplaçant la valeur par défaut `tez.task.relaxed.locality` de par false
Amélioration	Ajustez les configurations pour désactiver le délai dû à la localité et permettre une solution de repli non locale
Amélioration	TEZ-4547 : Ajoutez Tez AM JoBid au JobConf

Amazon EMR 7.5.0 - Notes de mise à jour de Tez

Amazon EMR 7.5.0 - Testez les modifications

Type	Description
Amélioration	Préemption de la tâche activée par défaut
Amélioration	Augmentation du nombre de lecteurs d'enregis trements permettant une initialisation asynchron e et proactive à 10
Amélioration	TEZ-4580 : Préemption lente des nouveaux conteneurs lorsque la réutilisation est activée

Amazon EMR 7.4.0 - Notes de mise à jour de Tez

Amazon EMR 7.4.0 - Testez les modifications

Type	Description
Amélioration	Supprimer l' EventCounter appendice de Tez supprimé dans Hadoop 3.4.0 via HADOOP-17 524
Upgrade	Mettre à jour commons-cli vers la version 1.5.0
Upgrade	Mettre à jour la version Bouncy-castle vers la version 1.70.0

Amazon EMR 7.2.0 - Notes de mise à jour de Tez

Amazon EMR 7.2.0 - Testez les modifications

Type	Description
Amélioration	Option fournie pour désactiver les contraintes de localité lors de la demande de conteneurs pour des tâches.
Correctif de bogue.	Rendre TaskDependencies# addTaskDependency thread sûr
Upgrade	Mettez à jour la version TLS vers la version 1.3 pour Tez.

Amazon EMR 7.2.0 - Nouvelles configurations

Classification	Nom	Par défaut	Description
tez-site	tez.task.relaxed.locality	false	Lorsque cette option est activée, les contraintes de localisation des racks et des nœuds ne sont pas prises en compte lors de la demande d'un conteneur pour une tâche.

Amazon EMR 7.1.0 - Notes de mise à jour de Tez

Amazon EMR 7.1.0 - Testez les modifications

Type	Description
Correctif de bogue.	TEZ-4394 — Netty4 ShuffleHandler : doit utiliser 1 thread de boss (#256)
Upgrade	Met à jour Netty vers la version 4.1.100.Final.
Upgrade	Met à jour Jetty vers la version 9.4.53.v2 0231009.

Amazon EMR 7.0.0 - Notes de mise à jour de Tez

Amazon EMR 7.0.0 Tez contient toutes les modifications et mises à jour apportées à Amazon EMR-6.15.0 Tez inclus.

Amazon EMR 7.0.0 - Tez les modifications

Type	Description
Upgrade	Tez Runtime utilise désormais Java 17 par défaut. Reportez-vous au guide de mise à jour d'EMR 7.0.0 pour plus de détails.

Amazon EMR 6.15.0 – Notes de mise à jour de Tez

Amazon EMR 6.15.0 – Modifications de Tez

Type	Description
Fonctionnalité	TEZ-4397 : permet d’ouvrir les divisions d’entrée Tez Input de manière asynchrone
Upgrade	TEZ-4493 : permet de mettre à niveau Apache Hadoop vers la version 3.3.6

Amazon EMR 6.15.0 – Fonctionnalités de Tez

- Ouverture de division asynchrone Tez – Amazon EMR 6.15.0 introduit des configurations que vous pouvez spécifier pour ouvrir de manière asynchrone les divisions d'entrée dans une division groupée Tez. La fonctionnalité a été initiée par TEZ-4397, mais a connu des régressions dans OSS Hive. Amazon EMR Hive a corrigé les régressions et les bogues supplémentaires dans la table Hive ACID. Cette amélioration permet d'accélérer les performances des requêtes de lecture lorsqu'il existe un grand nombre de divisions d'entrée dans une seule division groupée Tez. Pour plus d'informations, voir la rubrique Ouverture fractionnée asynchrone Tez.

Amazon EMR 6.14.0 - Notes de mise à jour de Tez

Amazon EMR 6.14.0 - Modifications de Tez

Type	Description
Amélioration	Mettre à jour la version TLS dans Tez vers la version 1.2

Amazon EMR 6.13.0 - Notes de mise à jour de Tez

Amazon EMR 6.13.0 - Tez les modifications

Type	Description
Correctif de bogue.	ANNULER TEZ-4295 : Impossible de décompresser les données. La longueur de la mémoire tampon est trop petite.
Correctif de bogue.	ANNULER TEZ-4302 : `NullPointerException` dans `CodecUtils` avec `GzipCodec` .
Correctif de bogue.	ANNULER TEZ-4234 : Le compresseur peut provoquer `IllegalArgumentException` dans `Buffer.limit` où la limite dépasse la capacité.

Type	Description
Correctif de bogue.	ANNULER TEZ-4135 : Améliorer l'allocation de mémoire lors de l'exécution de lectures en mémoire.

Amazon EMR 6.12.0 - Notes de mise à jour de Tez

Amazon EMR 6.12.0 - Tez les modifications

Type	Description
Amélioration	Ajout de la prise en charge de l'exécution JDK 11 et JDK 17
Correctif de bogue.	TEZ-4492 : Mettre à jour Bowerrc pour utiliser le miroir bower.herokuapp afin d'éviter le problème CERT_EXPIRE du registre Bower (BOWER-2608)
Upgrade	Mise à niveau de Surefire vers la version 3.0.0-M7

Amazon EMR 6.11.0 - Notes de mise à jour de Tez

Amazon EMR 6.11.0 - Tez les modifications

Type	Description
Bogue	Correction d'une transition d'état de sommet non valide lors du nettoyage des données réorganisées au niveau des sommets
Bogue	Le nettoyage fixe au niveau du DAG ou du sommet des données réorganisées ne fonctionne pas

Type	Description
Amélioration	Activez tez.am.dag.cleanup.on.completion par défaut pour effacer les données de shuffle une fois terminé DAGs

Amazon EMR 6.10.0 - Notes de mise à jour de Tez

Amazon EMR 6.10.0 - Modifications de Tez

Type	Description
Fonctionnalité	Activation de `tez.runtime.transfer.data-via-events.enabled` par défaut
Rétroportage	TEZ-4450 : Correction d'un échec de récupération des données de shuffle lorsque les données de shuffle sont transférées via des événements de déplacement de données
Rétroportage	TEZ-4460 : correction d'une erreur de délai de lecture lors de la récupération des données de shuffle depuis le gestionnaire Tez Shuffle
Rétroportage	TEZ-4455 : Ajout d'un ShuffleHandler pipeline pour une meilleure LoggingHandler débuggabilité
Bogue	Corrige le blocage de la tâche Tez par intermittence lorsque la préemption de la tâche est activée

Amazon EMR 6.9.0 - Notes de mise à jour de Tez

Amazon EMR 6.9.0 - Modifications de Tez

Type	Description
Upgrade	Tez est mis à niveau vers la version 0.10.2. Pour plus d'informations, consultez le journal des modifications pour Apache Tez 0.10.2.
Upgrade	Mise à niveau de Hadoop vers la version 3.3.3.
Bogue	Désactiver `tez.runtime.transfer.data-via-events.enabled` par défaut en raison du TEZ-4450.

Amazon EMR 6.8.0 - Notes de mise à jour de Tez

Amazon EMR 6.8.0 - Modifications de Tez

Type	Description
Rétroportage	TEZ-3363 : Supprimer les données intermédiaires au niveau du sommet pour Shuffle Handler
Rétroportage	TEZ-4129 : Supprimer les données de tentative intermédiaires en cas d'échec des tentatives pour Shuffle Handler
Rétroportage	TEZ-4430 : Correction de la propriété tez.task.launch.cmd-opts qui ne fonctionnait pas

Amazon EMR 6.7.0 - Notes de mise à jour de Tez

Amazon EMR 6.7.0 - Modifications de Tez

Type	Description
Rétroportage	TEZ-4403 : mise à niveau de la version SLF4 J vers la version 1.7.36
Rétroportage	TEZ-4405 : Remplacer log4j 1.x par reload4j
Rétroportage	TEZ-4411 : Échec de la compilation de Tez : .js introuvable FileSaver

Amazon EMR 6.6.0 - Notes de mise à jour de Tez

Amazon EMR 6.6.0 - Modifications de Tez

Type	Description
Rétroportage	TEZ-3918 : Correction de la propriété tez.task. log.level qui ne fonctionnait pas.
Rétroportage	TEZ-4353 : Mise à jour de commons-io vers la version 2.8.0.
Rétroportage	TEZ-4114 : Supprimer la dépendance directe de tez à la jetée.
Rétroportage	TEZ-4323 : Les jars Jetty ont été retirés du package de distribution avec TEZ-4114.

Trino

Trino est un moteur de requêtes open source conçu pour les requêtes interactives sur un large éventail de sources de données. Il peut s'agir de bases de données relationnelles, de données basées sur des fichiers, de données HDFS, etc. L'objectif le plus courant de Trino avec Amazon EMR est d'exécuter des requêtes SQL complexes sur de grands ensembles de données stockés dans Amazon S3. Il est également conforme à la norme ANSI SQL, ce qui le rend familier aux ingénieurs de bases de données, aux analystes de données et aux scientifiques des données familiarisés avec le SQL.

> **Note**
>
> PrestoSQL a été renommé Trino en décembre 2020. Les versions 6.4.0 et ultérieures d'Amazon EMR font généralement référence à Trino, tandis que les versions antérieures font référence à PrestoSQL.

> **Important**
>
> PrestoSQL, la version précédente de Trino, est toujours disponible pour une utilisation avec Amazon EMR. Cependant, nous recommandons vivement d'utiliser Trino à l'avenir avec Amazon EMR. Notez également que Trino et PrestoSQL ne peuvent pas fonctionner simultanément sur le même cluster.

Le tableau suivant répertorie la version de Trino incluse dans la dernière version d'Amazon EMR 7.x, ainsi que les composants qu'Amazon EMR installe avec Trino. Pour connaître la version des composants installés avec Trino dans cette version, consultez les versions des composants de la version 7.9.0.

Informations sur la version Trino

Étiquette de version Amazon EMR	Version Trino	Composants installés avec Trino
emr-7.9.0	Trino 467	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode,

Étiquette de version Amazon EMR	Version Trino	Composants installés avec Trino
		hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker

Rubriques

- Histoire et design de Trino
- Premiers pas avec Trino
- Configuration de Trino
- Bonnes pratiques pour Trino sur Amazon EMR
- Considérations relatives à Trino
- Historique des versions de Trino

Histoire et design de Trino

Trino est spécialisé dans l'interrogation de grands ensembles de données provenant de nombreuses sources différentes. Trino peut accéder à HDFS et l'interroger dans un cas d'utilisation traditionnel des mégadonnées, mais il peut également interroger des sources supplémentaires telles que des bases de données relationnelles et des bases de données NoSQL. Trino a débuté en tant que fork du moteur de requêtes Presto, en 2019. Depuis, il a été développé indépendamment de la base de code Presto.

Pour plus d'informations sur le moteur de requêtes Trino et son utilisation, consultez le site Web de Trino. Pour lire la documentation source de Trino, voir Présentation de Trino.

Concepts architecturaux

Trino peut exécuter des requêtes rapides et efficaces car il traite les données en parallèle au sein d'un cluster. Il a été conçu pour interroger un lac de données, car il est spécialisé pour les requêtes portant sur de gros volumes de données, généralement dans les cas d'utilisation impliquant Hadoop et HDFS. Mais il peut également interroger les bases de données relationnelles traditionnelles. Pour plus d'informations, consultez la section Architecture dans la documentation de Trino.

Composants de Trino

Trino possède quelques composants d'architecture clés qui fonctionnent ensemble pour accélérer l'exécution des requêtes. Il est utile d'avoir une connaissance pratique de ces éléments lorsque vous peaufinez votre cluster pour améliorer ses performances :

- Le coordinateur est responsable de l'orchestration des requêtes. Il analyse et optimise les requêtes SQL entrantes, génère des plans d'exécution, assigne des tâches aux nœuds de travail, collecte et assemble les résultats des requêtes. En outre, il surveille l'utilisation des ressources et suit l'état des nœuds de travail. Pour plus d'informations, consultez Coordinator dans la documentation de Trino.
- Les nœuds de travail gèrent le traitement des données pour les requêtes. Une fois que le coordinateur a attribué les tâches, les travailleurs récupèrent les données, effectuent les opérations nécessaires, telles que les jointures et les agrégations, et échangent des données intermédiaires avec d'autres travailleurs. Pour plus d'informations, consultez Worker dans la documentation de Trino.
- Les connecteurs sont des plug-ins qui permettent à Trino de se connecter à diverses sources de données et de les interroger. Chaque connecteur sait comment accéder aux données et les récupérer à partir de sa source, telle qu'Amazon S3, Apache Hive ou des bases de données relationnelles. Ces connecteurs mappent les données source à la structure du schéma de Trino.
- Un catalogue est un ensemble logique de schémas et de tables associés à un connecteur spécifique. Définis dans le coordinateur, les catalogues permettent à Trino de traiter différentes sources de données comme un seul espace de noms. Cela permet aux utilisateurs d'interroger plusieurs sources ensemble, telles que Hive et MySQL, de manière unifiée dans la même requête.
- Les clients tels que la CLI Trino se connectent au coordinateur Trino au moyen de pilotes JDBC et ODBC pour soumettre des requêtes SQL. Le coordinateur gère le cycle de vie des requêtes et fournit les résultats au client pour une analyse ou un reporting plus approfondis.

Exécution de requêtes

Pour comprendre comment Trino prend les instructions SQL et les exécute sous forme de requêtes, consultez les concepts de Trino dans la documentation de Trino.

Premiers pas avec Trino

Les procédures de cette section vous montrent comment configurer un cluster Amazon EMR afin d'interroger les sources de données du métastore avec Trino. Ces métastores, qui incluent le catalogue de données AWS Glue, stockent les métadonnées et les objets de base de données et gèrent les autorisations d'accès. Les procédures couvrent les prérequis, les paramètres de configuration recommandés, la création de connecteurs et l'exécution de requêtes sur des tables de métastore.

Rubriques

- Suivez les étapes préalables à l'utilisation d'Amazon EMR avec Trino
- Lancement d'un cluster Amazon EMR avec Trino
- Connexion au nœud primaire pour le cluster Amazon EMR et exécuter des requêtes

Suivez les étapes préalables à l'utilisation d'Amazon EMR avec Trino

Si vous n'avez pas utilisé AWS, ou si vous n'avez pas créé de cluster Amazon EMR, suivez ces étapes préalables avant de créer un cluster Amazon EMR avec Trino.

AWS configuration de l'environnement

Effectuez les étapes suivantes pour configurer votre AWS compte, si vous ne l'avez pas déjà fait :

1. Ouvrez un AWS compte, si vous n'en avez pas. Pour plus d'informations, consultez la section Création d'un AWS compte dans le Guide de référence AWS sur la gestion des comptes.
2. Connectez-vous à votre compte en tant que l'utilisateur administratif.
3. Créez un groupe et attribuez-lui des utilisateurs.
4. Créez une paire de EC2 clés Amazon, que vous pourrez utiliser ultérieurement pour sécuriser la communication entre les ressources avec SSH. Cette étape est requise si vous prévoyez de vous connecter au nœud primaire pour effectuer des tâches. Pour plus d'informations, consultez Connexion au nœud primaire du cluster Amazon EMR à l'aide de SSH.

Lancement d'un cluster Amazon EMR avec Trino

Voici une description des choix de configuration corrects lorsque vous créez un cluster avec Trino.

Utilisation d'un connecteur Hive pour rendre les données disponibles pour les requêtes

Vous pouvez configurer un connecteur Trino pour un métastore Hive afin d'interroger les données du métastore provenant de votre cluster. Un métastore est une couche d'abstraction qui rend le contenu ou les données basés sur des fichiers disponibles sous forme de tables, ce qui facilite les requêtes. Vous devez configurer un connecteur dans Amazon EMR pour mettre les tables de métastore Hive à la disposition du cluster. Il suffit de procéder comme indiqué ci-dessous :

1. Choisissez AWS Glue dans la console et créez un tableau en fonction de vos données sources dans Amazon S3. Un tableau du catalogue de données AWS Glue est la définition des métadonnées des données. Dans ce contexte, il est judicieux de créer le tableau manuellement, en créant des colonnes comme vous le souhaitez, à partir de vos données sources. Pour plus d'informations sur la création de tables dans AWS Glue à partir de données semi-structurées dans Amazon S3, consultez la section Création de tables à l'aide de la console dans le guide de l'utilisateur de AWS Glue.
2. Définissez votre configuration dans le cadre de la création du cluster. Sélectionnez l'onglet Configurations. Les configurations sont des spécifications facultatives pour votre cluster. Lorsque vous entrez une configuration, ajoutez du JSON comme dans l'exemple suivant, qui indique à Trino d'utiliser le catalogue de données AWS Glue comme métastore Hive externe pour les métadonnées des tables :

   ```
   {
       "classification": "trino-connector-hive",
       "properties": {
           "hive.metastore": "glue"
       }
   }
   ```

   Vous pouvez également appliquer des configurations dans la section Paramètres du logiciel lorsque vous créez un cluster.

   En outre, vous pouvez configurer d'autres types de connecteurs, par exemple pour vous connecter à Apache Iceberg. Pour plus d'informations, consultez la section Utiliser un cluster Iceberg avec Trino dans le guide de mise à jour d'Amazon EMR. La configuration de paramètres supplémentaires est facultative.

Pour poursuivre les étapes de démarrage, voir. Connexion au nœud primaire pour le cluster Amazon EMR et exécuter des requêtes

Création d'un cluster avec Trino

Ce qui suit décrit les choix de configuration corrects lorsque vous créez un cluster que vous souhaitez utiliser avec Trino.

> **Important**
>
> Avant de créer votre cluster, complétez AWS la configuration de Glue Data Catalog en tant que métastore Hive, ce que nous recommandons pour démarrer. Pour de plus amples informations, veuillez consulter Utilisation d'un connecteur Hive pour rendre les données disponibles pour les requêtes.

1. Dans la AWS console, sélectionnez Amazon EMR dans les services. Lorsque vous choisissez Amazon EMR, si vous avez des clusters existants, vos EMR sur EC2 les clusters sont répertoriés.
2. Choisissez Créer un cluster. À partir de là, vous lancez le processus de création d'un cluster.
3. Donnez un nom à votre cluster et choisissez une version d'Amazon EMR. Vous pouvez choisir la version la plus récente pour le didacticiel.
4. Choisissez le pack Trino dans lequel l'application Trino est présélectionnée. Les offres groupées sont configurées pour des raisons de commodité lorsque vous connaissez à l'avance l'objectif du cluster. Sinon, cochez simplement la case Trino.
5. Pour la configuration du cluster, choisissez Uniform instance groups. Allez-y et supprimez des groupes d'instances supplémentaires.
6. Choisissez un type d'instance. En général, nous vous recommandons de choisir un type d'instance avec au moins 16 GiB de mémoire. De plus, pour le dimensionnement et le provisionnement du cluster, choisissez Définir la taille du cluster manuellement.
7. À ce stade, définissez la configuration de votre métastore Hive pour qu'elle pointe vers Glue. AWS Ceci est détaillé dans la sectionUtilisation d'un connecteur Hive pour rendre les données disponibles pour les requêtes. Effectuez cette opération avant de créer le cluster.
8. Choisissez Créer un cluster. La fin du processus peut prendre quelques minutes.

Les étapes décrites ici ne couvrent pas toutes les étapes de configuration en détail. Plus d'informations sur la configuration d'un cluster sont disponibles sur Planifier, configurer et lancer des clusters Amazon EMR.

> **Note**
>
> Ne sélectionnez pas Presto et Trino pour une utilisation sur le même cluster. Il n'est pas pris en charge de les exécuter ensemble. Il est également recommandé que si vous exécutez Trino, vous n'exécutiez aucune autre application sur le cluster, telle que Spark.

Connexion au nœud primaire pour le cluster Amazon EMR et exécuter des requêtes

Fournir des données de test et configurer les autorisations

Vous pouvez tester Amazon EMR avec Trino en utilisant AWS Glue Data Catalog et son métastore Hive. Les étapes préalables suivantes décrivent comment configurer les données de test, si vous ne l'avez pas déjà fait :

1. Créez une clé SSH à utiliser pour le chiffrement des communications, si vous ne l'avez pas déjà fait.
2. Vous pouvez choisir parmi plusieurs systèmes de fichiers pour stocker les données et les fichiers journaux. Pour commencer, créez un compartiment Amazon S3. Donnez un nom unique au compartiment. Lorsque vous le créez, spécifiez la clé de chiffrement que vous avez créée.

   > **Note**
   >
   > Choisissez la même région pour créer à la fois votre compartiment de stockage et le cluster Amazon EMR.

3. Choisissez le bucket que vous avez créé. Choisissez Créer un dossier et attribuez au dossier un nom mémorable. Lorsque vous créez le dossier, choisissez une configuration de sécurité. Vous pouvez choisir les paramètres de sécurité pour le parent ou les personnaliser davantage.

4. Ajoutez des données de test à votre dossier. Pour les besoins de ce didacticiel, l'utilisation d'un fichier .csv composé d'enregistrements séparés par des virgules fonctionne bien pour compléter ce cas d'utilisation.
5. Après avoir ajouté des données dans un compartiment Amazon S3, configurez une table dans AWS Glue pour fournir une couche d'abstraction permettant d'interroger les données.

Connect et exécution de requêtes

Ce qui suit décrit comment vous vous connectez à un cluster exécutant Trino et comment exécuter des requêtes sur celui-ci. Avant cela, assurez-vous de configurer le connecteur de métastore Hive, décrit dans la procédure précédente, afin que les tables de métastore soient visibles.

1. Nous vous recommandons d'utiliser EC2 Instance Connect pour vous connecter à votre cluster, car il fournit une connexion sécurisée. Choisissez Connect to the primary node using SSH dans le résumé du cluster. La connexion nécessite que le groupe de sécurité dispose d'une règle entrante pour autoriser les connexions via le port 22 aux clients du sous-réseau. Vous devez également utiliser l'utilisateur hadoop lors de la connexion.
2. Démarrez la CLI Trino en exécutant. `trino-cli` Cela vous permet d'exécuter des commandes et d'interroger des données avec Trino.
3. Exécutez `show catalogs;`. Vérifiez que le catalogue des ruches est répertorié. Cela fournit une liste des catalogues disponibles, qui contiennent des magasins de données ou des paramètres système.
4. Pour voir les schémas disponibles, exécutez`show schemas in hive;`. À partir de là, vous pouvez exécuter `use` *`schema-name`*`;` et inclure le nom de votre schéma. Ensuite, vous pouvez courir `show tables;` pour répertorier les tables.
5. Interrogez une table en exécutant une commande telle que`SELECT * FROM` *`table-name`*, en utilisant le nom d'une table dans votre schéma. Si vous avez déjà exécuté l'`USE`instruction pour vous connecter à un schéma spécifique, il n'est pas nécessaire d'utiliser une notation en deux parties telle que*`schema`*. *`table`*.

Configuration de Trino

Rubriques

- Configuration des connecteurs pour Trino
- Surveillance

Configuration des connecteurs pour Trino

Connexion à AWS Glue en tant que métastore Hive

Il est important et utile de comprendre que vous pouvez configurer AWS Glue Data Catalog comme métastore Hive lorsque vous exécutez des requêtes avec Trino. Pour plus d'informations, y compris les étapes de configuration d'un cluster avec un métastore Hive, voir Utilisation du catalogue de données AWS Glue comme métastore pour Hive.

Pour plus d'informations sur l'intégration d'EMR sur EKS avec AWS Glue, consultez les meilleures pratiques suivantes, Intégration des conteneurs EMR à Glue. AWS

Connexion aux tables Iceberg lors de l'utilisation de Trino avec Amazon EMR

Iceberg est un format de table ouvert pour les tableaux analytiques. Il a été créé pour des moteurs tels que Spark et Trino afin d'interroger des données volumineuses à partir des mêmes tables, à l'aide de requêtes SQL. Il inclut des fonctionnalités telles que l'isolation des lectures et des écritures de données, afin qu'un lecteur puisse éviter d'interroger des données partiellement mises à jour, par exemple. Il prend également en charge les fonctionnalités d'état, telles que les instantanés. Il fournit une couche d'abstraction grâce à l'utilisation de métadonnées et de fichiers manifestes. Ils décrivent le schéma des tables et facilitent les requêtes de données sans avoir à connaître beaucoup de détails sur leur mise en forme ou leur organisation. Lorsque vous êtes connecté, vous pouvez à la fois lire les données des tables, les mettre à jour ou écrire de nouvelles données dans les fichiers sous-jacents.

Un atelier vous montre comment configurer les tables Iceberg avec Amazon EMR et AWS Glue. Pour plus d'informations, consultez Analytics Workshop - Configurer et utiliser les tables Apache Iceberg sur votre lac de données.

Établir des liens avec les clients

Vous pouvez vous connecter à Trino à l'aide d'un pilote JDBC disponible. Pour plus d'informations, consultez le pilote JDBC dans la documentation de Trino.

Surveillance

Vous pouvez surveiller les clusters Amazon EMR via le. AWS Management Console Pour plus d'informations, consultez Afficher et surveiller un cluster Amazon EMR pendant qu'il exécute des tâches. Amazon EMR envoie également ses mesures de surveillance à. Amazon CloudWatch Pour

plus d'informations sur la surveillance d'un cluster Amazon EMR, consultez les Amazon CloudWatch événements et mesures Amazon EMR.

Bonnes pratiques pour Trino sur Amazon EMR

L'architecture de Trino est conçue pour des requêtes SQL rapides et distribuées sur de grands ensembles de données provenant de plusieurs sources de données, selon un modèle coordinateur-travailleur, dans lequel chaque composant joue un rôle spécialisé dans l'exécution des requêtes. Il existe quelques domaines ou catégories sur lesquels vous pouvez vous concentrer afin de configurer votre cluster Amazon EMR exécutant Trino pour obtenir les meilleures performances. Tel est le cas des éléments suivants :

- Ajustement des paramètres de configuration du cluster pour optimiser la mémoire.
- Optimisation des paramètres de partitionnement et de distribution des données.
- Utilisation du filtrage dynamique pour réduire le nombre de résultats des requêtes.

Certains de ces paramètres sont réglés automatiquement lorsque vous utilisez Trino avec Amazon EMR. D'autres peuvent être définis manuellement via la console ou via les commandes CLI. Les rubriques de cette section vous aident à configurer vos données et votre cluster de manière optimale.

Rubriques

- Principaux domaines d'intérêt pour l'amélioration des performances
- Collecter et utiliser les statistiques des tables
- Défis courants liés à l'évolutivité des charges de travail de Trino

Principaux domaines d'intérêt pour l'amélioration des performances

Trino maximise le parallélisme des requêtes et l'optimisation de la mémoire. Cette architecture apporte de la flexibilité en lui permettant d'interroger des sources de données multiples et variées tout en assurant une mise à l'échelle efficace. Les domaines clés d'amélioration des performances de Trino sont ceux répertoriés ci-dessous.

Optimisation de la mémoire

La gestion de la mémoire dans Trino est essentielle pour atteindre des performances et une stabilité élevées, en particulier lorsque vous exécutez des requêtes complexes et volumineuses. Trino utilise un modèle de mémoire distribuée. Dans ce modèle, la mémoire est allouée entre les nœuds de

travail pour le traitement des tâches, des agrégations, des jointures et d'autres opérations. La liste suivante présente un ensemble de ces paramètres :

- query.max-memory — Définit la mémoire maximale disponible pour une seule requête sur l'ensemble du cluster. Il s'agit d'une limite stricte ; si une requête dépasse cette mémoire, elle échouera.
- requête. max-memory-per-node — Définit la mémoire maximale qu'une requête peut consommer sur chaque nœud de travail. Cette configuration garantit qu'aucune requête ne monopolise les ressources d'un travailleur.
- Taille du tas de mémoire de machine virtuelle Java : configuré au niveau de la machine virtuelle Java, il définit la taille de segment de mémoire maximale pour le processus du serveur Trino sur chaque nœud. Cette valeur doit généralement être supérieure aux configurations liées à la mémoire (il s'agit de la somme des requêtes). max-memory-per-nodeet de la mémoire. heap-headroom-per-node) dans Trino pour éviter que le système ne manque de mémoire au niveau de la JVM.
- mémoire. heap-headroom-per-node — Spécifie la quantité de mémoire tampon à exclure de la taille du segment de mémoire JVM pour les opérations autres que les requêtes. Cela est crucial pour garantir des frais généraux suffisants pour les opérations internes et la collecte des déchets.

Filtrage dynamique

Le filtrage dynamique dans Trino est une technique d'optimisation qui améliore les performances des requêtes en réduisant la quantité de données traitées, en particulier lors des jointures. Il applique dynamiquement des conditions de filtre pour limiter les données scannées d'un côté d'une jointure, en fonction des données visibles de l'autre côté, ce qui est particulièrement utile dans les requêtes où un côté de la jointure est très sélectif (c'est-à-dire qu'il contient un petit sous-ensemble de données). Il est activé par défaut sur Amazon EMR. Voici un exemple de requête :

```
SELECT orders.order_id, orders.total_amount
FROM orders
JOIN customers ON orders.customer_id = customers.customer_id
WHERE customers.country = 'France';
```

Sans filtrage dynamique, Trino analyse l'intégralité du tableau des commandes dans une jointure, même si seul un petit sous-ensemble de clients (ceux de France) est pertinent. Cette approche permet de lire toutes les lignes du tableau des commandes, ce qui entraîne des coûts d'E/S et de traitement élevés. Avec le filtrage dynamique, Trino analyse d'abord la petite table des clients,

récupère les valeurs customer_id uniquement pour les clients de France, puis applique ce sous-ensemble en tant que filtre sur les commandes. Cela signifie que seules les lignes pertinentes des commandes (celles dont le customer_id correspond au sous-ensemble filtré) sont scannées, ce qui réduit considérablement le nombre d'enregistrements traités.

Spill to Disk

Dans Trino, le déversement de disque permet de transférer les résultats des requêtes intermédiaires sur le disque, ce qui permet de terminer les requêtes gourmandes en mémoire, même si elles dépassent les limites de mémoire définies par ou. `query_max_memory` `query_max_memory_per_node` Par défaut, Trino applique ces limites pour garantir une allocation de mémoire équitable et empêcher le blocage du cluster. Toutefois, lorsqu'une requête volumineuse dépasse ces limites, elle risque d'être interrompue. Le déversement de disque résout ce problème en utilisant`revocable memory`, permettant à une requête d'emprunter de la mémoire supplémentaire qui peut être révoquée si des ressources sont nécessaires ailleurs. Lorsque la mémoire est révoquée, les données intermédiaires se répandent sur le disque, ce qui permet aux requêtes de poursuivre le traitement sans dépasser les limites de mémoire. Veuillez noter qu'une requête qui est forcée de se propager sur le disque peut avoir un temps d'exécution plus long, c'est pourquoi elle est désactivée par défaut. Pour activer Spill sur Amazon EMR, utilisez la configuration suivante :

- `spill-enabled=true`— Permet de renverser le disque lorsque l'utilisation de la mémoire dépasse les seuils disponibles.
- `spill-paths`— Définit les répertoires dans lesquels les données déversées sont stockées, `spill-paths=/mnt/spill`

Collecter et utiliser les statistiques des tables

La collecte des statistiques des tables permet à l'optimiseur basé sur les coûts de Trino de prendre des décisions éclairées concernant les ordres de jointure, le filtrage et l'élagage des partitions, ce qui se traduit par de meilleures performances.

Vous pouvez utiliser la `ANALYZE` commande pour collecter des statistiques pour les tables Hive ou Iceberg :

```
ANALYZE sales;
```

La collecte de statistiques sur de larges tableaux peut être une lourde charge pour les ressources. Nous vous recommandons de spécifier un sous-ensemble de colonnes utilisées dans les jointures, les filtres ou les opérations de regroupement.

Il s'agit d'une autre commande utile. Il affiche les statistiques actuelles d'un tableau afin de vérifier si les statistiques sont à jour.

```
show stats for table_name;
```

Défis courants liés à l'évolutivité des charges de travail de Trino

Les principaux avantages de l'utilisation d'Amazon S3 avec Trino sont sa capacité à s'adapter à de grands volumes de données et sa rentabilité. Mais lorsque vous interrogez de gros volumes de données, une série de problèmes de performances connexes peuvent parfois survenir. Cela peut être dû à la manière dont les données sont stockées, à des paramètres de configuration qui limitent les bonnes performances ou à d'autres raisons. Lorsque ces problèmes surviennent, vous pouvez prendre des mesures efficaces pour les éviter ou les atténuer.

Cette section commence par une liste d'optimisations générales que vous pouvez mettre en œuvre pour améliorer les performances des requêtes sur de grands volumes de données. Ensuite, les problèmes courants sont détaillés et des mesures d'atténuation sont proposées pour chacun d'entre eux.

Ce sujet est tiré de la présentation de conférence suivante : Accélérer les performances à grande échelle : meilleures pratiques pour Trino avec Amazon S3.

Optimisation de la mise en page des données volumineux

Les problèmes de performances ne sont pas rares lorsque vous interrogez de grands ensembles de données. Mais il existe des bonnes pratiques que vous pouvez mettre en œuvre pour avoir une longueur d'avance lorsque vous utilisez Trino pour interroger des données dans Amazon S3. Tel est le cas des éléments suivants :

- Partitionnement — Le partitionnement consiste à organiser les données dans une hiérarchie et à stocker les données associées ensemble, en fonction d'attributs connexes. Le partitionnement permet aux requêtes de ne pas avoir à analyser autant de données non pertinentes et d'améliorer les performances des requêtes. Vous pouvez utiliser différentes stratégies de partitionnement, telles que l'organisation des données sources à l'aide de préfixes, notamment par plages de dates, régions ou autres attributs. Pour plus d'informations sur le partitionnement des données dans

Amazon S3 afin d'améliorer les performances, consultez le billet de blog Get started managing partitions for Amazon S3 tables backed by the AWS Glue Data Catalog ou le billet Top 10 des conseils d'optimisation des performances pour Amazon Athena.

- Compartitionnement — Le découpage regroupe des données connexes dans des fichiers communs. Par exemple, si vous interrogez des données en fonction d'une région géographique, comme un état, vous pouvez améliorer les performances des requêtes en regroupant toutes les données relatives à un état donné dans le même fichier ou groupe de fichiers. Pour que cela fonctionne au mieux, basez votre bucket sur un attribut de données à forte cardinalité, tel qu'un État ou une province, par exemple. Vous pouvez également prendre en compte vos modèles de requêtes. Par exemple, il peut s'agir de regrouper les données de la Californie et de l'Oregon, si vos requêtes lisent généralement les données de ces États ensemble.
- Gestion des préfixes S3 — Vous pouvez utiliser les préfixes Amazon S3 pour implémenter une stratégie de partitionnement. Si vous n'utilisez qu'un seul préfixe pour un compartiment Amazon S3, comme une date précise, par exemple, cela peut entraîner un nombre élevé de demandes et une erreur HTTP 503. Nous vous recommandons d'utiliser des préfixes pour ajouter des conditions supplémentaires et organiser vos données sources de manière plus efficace. Pour plus d'informations, consultez Organisation des objets à l'aide de préfixes dans la documentation Amazon S3. Le bref exemple suivant montre un préfixe qui améliore le débit des demandes :. `s3://bucket/country=US/dt=2024-06-13` Dans cet exemple, le pays et la date sont inclus dans le préfixe, ce qui entraîne moins de lectures que dans le cas où le préfixe inclut uniquement la date.

  La réduction des erreurs HTTP 503 est abordée plus en détail dans la section sur le ralentissement du protocole HTTP qui suit dans cette rubrique.
- Optimisation de la taille des données : vous pouvez exécuter la commande OPTIMIZE pour définir une configuration favorable aux requêtes les plus performantes. Pour l'exécuter sur des tables externes Hive, procédez comme suit :
  - À utiliser `OPTIMIZE` avec le paramètre suivant : `hive.non-managed-table-writes-enabled=true` Pour plus d'informations sur cette propriété, consultez la section Propriétés de configuration générales de Hive.
  - Définissez le paramètre de session suivant : `SET SESSION` *`catalog`*`.non_transactional_optimize_enabled=true`
  - Exécutez la `OPTIMIZE` commande :`ALTER TABLE` *`catalog`*`.`*`schema`*`.`*`table`* `EXECUTE optimize(file_size_threshold => '128MB')`. Dans ce cas, `file_size_threshold` c'est 100 Mo par défaut. L'augmentation de ce seuil, comme indiqué dans l'exemple, entraînera la fusion des fichiers inférieurs à 128 Mo.

- Configurer les tentatives — Vous pouvez augmenter la limite de tentatives, ce qui peut atténuer le risque d'erreurs HTTP 503, en définissant les paramètres suivants :. `s3.max-error-retries` Cela s'applique lorsque vous utilisez l' TrinoFileSystem API et la version Trino 449 ou ultérieure. D'autre part, dans le cas où vous utilisez Amazon EMR avec Trino, vous utilisez EMRFS pour accéder à Amazon S3. Avec EMRFS, vous pouvez augmenter le nombre de départs à la retraite en modifiant le paramètre. `fs.s3.maxRetries`
- Choisissez une classe de stockage Amazon S3 : le choix de la classe de stockage appropriée pour les données à différents stades de leur cycle de vie peut améliorer les performances et les coûts, en fonction de vos besoins en matière de collecte de données spécifiques. Pour plus d'informations, consultez Comprendre et gérer les classes de stockage Amazon S3 dans la documentation Amazon S3.
- Migrer vers Iceberg — Une autre solution pour atténuer les problèmes de performances, notamment en ce qui concerne l'exécution de requêtes sur de petits fichiers, consiste à migrer vers les tables Iceberg. Iceberg possède des fonctionnalités qui gèrent bien les petits fichiers.
- Utiliser le compactage automatique des données : si vous utilisez des tables Iceberg, le compactage automatique des données avec le catalogue de données AWS Glue peut optimiser la taille des données et améliorer les performances des requêtes.

Défis courants lorsque vous interrogez de grands ensembles de données

Cette section répertorie un ensemble de problèmes courants qui peuvent survenir lorsque vous accumulez un ensemble de données volumineux dans Amazon S3 et que vous l'interrogez avec Trino. Chaque section indique les moyens de résoudre le problème ou de réduire son impact sur les requêtes. Chacun des problèmes décrits dans les sections suivantes a été reproduit et testé à l'aide d'un connecteur Hive.

Analyses de données volumineuses

Lorsque votre requête doit analyser de grands ensembles de données, cela peut entraîner des problèmes tels que des performances de requête lentes et des coûts de stockage plus élevés. D'importants volumes de données peuvent résulter d'une croissance rapide des données ou d'une planification qui n'entraîne pas le déplacement des données existantes dans un délai approprié. Cela peut entraîner des requêtes plus lentes.

Pour limiter les pertes de performances liées à l'analyse de grands ensembles de données, nous vous recommandons d'utiliser le partitionnement et le partitionnement :

- Le partitionnement regroupe les données associées en fonction de leurs attributs. L'utilisation efficace du partitionnement peut améliorer considérablement les performances des requêtes.
- Le découpage consiste à regrouper des données dans des fichiers ou des compartiments en fonction de colonnes de données spécifiques et connexes. Le cloisonnement consiste généralement à conserver physiquement ensemble les fichiers de données sources connexes.

Pour illustrer comment l'atténuation peut fonctionner pour les analyses de données volumineuses, supposons que vous stockez et interrogez des données contenant des enregistrements dotés d'un attribut d'état, qui peut être attribué à la Californie ou à l'Alaska, et que cet attribut d'état est l'une des conditions de votre requête. Vous pouvez améliorer les performances des requêtes en stockant les données pour chaque état dans un compartiment S3 distinct ou en partitionnant vos données en fonction de l'état à l'aide d'un préfixe S3. Ce partitionnement et ce découpage peuvent également améliorer les performances si vous les basez sur une colonne supplémentaire, comme un attribut de date, par exemple.

> **Note**
>
> Si une colonne présente une cardinalité élevée et que vous souhaitez l'utiliser pour regrouper des données, nous vous recommandons d'utiliser le découpage dans ce cas. D'un autre côté, les clés de partition doivent généralement avoir une cardinalité inférieure.

Utilisation de différents types de stockage S3

En règle générale, vous choisissez les types de stockage en fonction des exigences de performance, d'accès aux données, de résilience et de coût pour vos charges de travail. Il peut y avoir des compromis entre le coût et les performances. Il est important de choisir la classe de stockage Amazon S3 appropriée qui correspond à vos modèles d'accès aux données. Il existe deux types d'accès principaux :

- Données accessibles de manière connue ou prévisible. En général, si vos données sont rarement consultées, S3 Standard IA peut être un bon choix, car il permet de réduire les coûts. Si vous avez fréquemment accédé à des données, S3 Standard est idéal pour l'accès avec Amazon EMR et Trino.
- Données auxquelles on accède de manière inconnue ou imprévisible. Cela peut nécessiter l'utilisation d'autres classes de stockage Amazon S3. Il existe des compromis entre les classes de stockage S3. Il s'agit notamment de la latence, du coût du stockage et de la disponibilité. Vous

pouvez choisir un type de stockage S3 approprié, en fonction de vos charges de travail et de vos modèles d'accès. Pour une description des avantages de chaque classe, consultez Classes de stockage Amazon S3.

Utilisation du compactage

Vous pouvez également utiliser le compactage automatique d'Iceberg, si vous utilisez des tables Iceberg, ce qui permet d'optimiser la taille des fichiers, afin d'améliorer l'efficacité des requêtes. Pour plus d'informations, consultez AWS Glue Data Catalog qui prend désormais en charge le compactage automatique des tables Apache Iceberg.

Erreurs de ralentissement du protocole HTTP

Cela se produit lorsque le taux de demandes dépasse un seuil préconfiguré sur un préfixe Amazon S3. L'erreur HTTP qui se produit le plus souvent lorsque cet état est atteint est la suivante : Erreur 503 : veuillez réduire votre taux de requêtes. La source de ce problème peut être liée à la présence d'un grand nombre de petits fichiers, en raison du nombre de divisions qui doivent être créées pour pouvoir lire les données. Il existe deux moyens d'atténuer le problème :

- Augmentez la limite de nouvelles tentatives pour les requêtes Amazon S3 dans Trino. Ceci est défini pour l'utilisation d'EMRFS `fs.s3.maxretries` dans Trino 449.
- Optimisez la taille des fichiers, ce qui peut également entraîner une baisse du taux de demandes.

Pour plus d'informations sur la façon dont Trino détermine le nombre de divisions dans un ensemble de données à interroger, consultez la section Propriétés de configuration du réglage des performances dans la documentation du connecteur Hive.

Difficulté à interroger les petits fichiers

L'interrogation de nombreux petits fichiers peut entraîner une surcharge d'E/S importante, en raison du nombre élevé de requêtes GET et LIST, et par conséquent affecter négativement les performances des requêtes. L'optimisation de la taille des fichiers peut améliorer les performances des requêtes. Il y a plusieurs manières de le faire :

- Consolidez les données dans des fichiers moins volumineux. (En général, nous recommandons de limiter la taille des fichiers à environ 128 Mo.) Vous pouvez le faire à l'aide d'outils lorsque vous ingérez des données, par exemple dans un pipeline ETL, ou vous pouvez consolider les données manuellement. Si vous ne disposez pas de ces solutions, les autres options vous conviennent peut-être mieux.

- Exécutez la commande `OPTIMIZE`.
- Définissez le paramètre `SESSION`.

Notez qu'Iceberg dispose d'une fonctionnalité permettant de fusionner de petits fichiers en fichiers plus volumineux, à savoir le compactage automatique. Il fonctionne avec les fichiers gérés avec le AWS Glue Data Catalog. Pour plus d'informations, consultez AWS Glue Data Catalog qui prend désormais en charge le compactage automatique des tables Apache Iceberg.

Requêtes contenant des données inutiles

Comme les données augmentent régulièrement, il est impératif de suivre vos habitudes d'accès aux données et de déplacer les données de manière appropriée au fur et à mesure qu'elles vieillissent ou deviennent inutiles. En effet, à mesure que les données augmentent, les performances des requêtes peuvent se dégrader au fil du temps, principalement en raison du volume considérable de données à analyser lors de l'exécution d'une requête. Amazon S3 et d'autres services proposent des conseils pour la migration du cycle de vie des données, qui présentent des stratégies pour déplacer les données vers différents sites de stockage lorsqu'il fait froid. Cela présente également un avantage en termes de coûts de stockage.

Outre la migration des données, vous pouvez utiliser d'autres stratégies, telles que la suppression des données sources qui ne sont pas pertinentes pour les requêtes que vous exécutez. Cela peut demander du travail, car cela peut impliquer de modifier votre schéma de données sources. Mais son résultat positif est de réduire le volume de données et d'accélérer les requêtes. Pour plus d'informations, consultez Gestion du cycle de vie des objets.

Considérations relatives à Trino

Tenez compte des facteurs suivants lorsque vous exécutez Trino EMR.

Propriétés de déploiement Trino non configurables

Le tableau suivant affiche les différentes options de configuration pour les `properties` fichiers Trino.

Fichier	Configurable
`log.properties`	Trino : configurable dans les versions 6.1.0 et ultérieures d'Amazon EMR. Utilisez la classific

Fichier	Configurable
	ation de configuration `prestosql-log` ou `trino-log` .
`config.properties`	Trino : configurable dans les versions 6.1.0 et ultérieures d'Amazon EMR. Utilisez la classific ation de configuration `prestosql-config` ou `trino-config` .
`hive.properties`	Trino : configurable dans les versions 6.1.0 et ultérieures d'Amazon EMR. Utilisez la classification de configuration `prestosql-connector-hive` ou `trino-connector-hive` .
`node.properties`	Trino : configurable dans les versions 6.1.0 et ultérieures d'Amazon EMR. Utilisez la classific ation de configuration `prestosql-node` ou `trino-node` .
`jvm.config`	Non configurable.

Considérations supplémentaires

- Pour Trino on EMR version 6.1.0 et versions ultérieures, Amazon EMR configure automatiquement une clé secrète partagée pour une communication interne sécurisée entre les nœuds du cluster. Vous n'avez pas besoin d'effectuer de configuration supplémentaire pour activer cette fonctionnalité de sécurité, et vous pouvez remplacer la configuration par votre propre clé secrète. Pour plus d'informations sur l'authentification interne de Trino, consultez la documentation Trino 353 : Communication interne sécurisée.

Historique des versions de Trino

Les sections des notes de publication détaillent les modifications et les mises à jour pour une version spécifique de Trino sur Amazon EMR.

Notes de mise à jour de Trino par version

- Amazon EMR 7.6.0
- Amazon EMR 7.3.0 - Notes de mise à jour de Trino
- Amazon EMR 6.9.0 - Notes de mise à jour de Trino

Amazon EMR 7.6.0

Amazon EMR 7.6.0

- Pour prendre en charge les requêtes de longue durée, Trino inclut désormais un mécanisme d'exécution tolérant aux pannes. L'exécution tolérante aux pannes atténue les échecs des requêtes en réessayant les requêtes qui ont échoué ou les tâches correspondantes.

Amazon EMR 7.6.0

Amazon EMR 7.6.0

Type	Description
Upgrade	Mise à niveau de Trino vers 457

Amazon EMR 7.3.0 - Notes de mise à jour de Trino

Amazon EMR 7.3.0

- Cette version met à niveau Trino de la version 436 à la version 442.
- Cette version redirige les requêtes Hudi vers le nouveau correcteur Hudi. L'ancien connecteur Hive ne peut plus lire les tables Hudi. Remarque
- Cette version supprime le module Rubix d'Amazon EMR car il est désormais obsolète en tant que logiciel libre.
- Cette version supprime le mode ancien dans la `hive.security` propriété. La valeur par défaut est maintenant`allow-all`.

Amazon EMR 6.9.0 - Notes de mise à jour de Trino

Amazon EMR 6.9.0 - Nouvelles fonctionnalités de Trino

- Pour prendre en charge les requêtes de longue durée, Trino inclut désormais un mécanisme d'exécution tolérant aux pannes. L'exécution tolérante aux pannes atténue les échecs des requêtes en réessayant les requêtes qui ont échoué ou les tâches correspondantes.

Amazon EMR 6.9.0 - Modifications apportées à Trino

Amazon EMR 6.9.0 - Modifications apportées à Trino

Type	Description
Upgrade	Mise à niveau de Trino vers la version 398
Upgrade	Prise en charge pour Hadoop 3.3.3.
Fonctionnalité	Support tardigrade : ajout de la prise en charge du spoulage des échanges sur HDFS et Amazon S3.
Correctif de bogue	Lorsque Trino Iceberg est utilisé et que le catalogue Glue est activé, évitez d'ajouter un URI de métastore dans `iceberg.properties`.

Amazon EMR 6.9.0 - Problèmes connus liés à Trino

- Pour Amazon EMR version 6.9.0, Trino ne fonctionne pas sur les clusters activés pour Apache Ranger. Si vous devez utiliser Trino avec Ranger, contactez Support.

Apache Zeppelin

Utilisez Apache Zeppelin sous forme de bloc-notes pour l'exploration des données interactives. Pour plus d'informations sur Zeppelin, consultez https://zeppelin.apache.org/. Zeppelin est inclus dans les versions Amazon EMR 5.0.0 et supérieures. Les versions antérieures incluent Zeppelin en tant qu'application sandbox. Pour de plus amples informations, veuillez consulter Versions Amazon EMR 4.x.

Pour accéder à l'interface web Zeppelin, configurez un tunnel SSH sur le nœud maître et une connexion proxy. Pour plus d'informations, consultez Affichage des interfaces web hébergées sur des clusters Amazon EMR.

Le tableau suivant répertorie la version de Zeppelin incluse dans la dernière version de la série Amazon EMR 7.x, ainsi que les composants qu'Amazon EMR installe avec Zeppelin.

Pour connaître la version des composants installés avec Zeppelin dans cette version, consultez Versions des composants de la version 7.9.0.

Informations sur la version de Zeppelin pour emr-7.9.0

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-7,9.0	Zeppelin 0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Le tableau suivant répertorie la version de Zeppelin incluse dans la dernière version de la série Amazon EMR 6.x, ainsi que les composants qu'Amazon EMR installe avec Zeppelin.

Pour connaître la version des composants installés avec Zeppelin dans cette version, consultez les versions des composants de la version 6.15.0.

Informations sur la version de Zeppelin pour emr-6.15.0

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.15.0	Zeppelin 0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Le tableau suivant répertorie la version de Zeppelin incluse dans la dernière version de la série Amazon EMR 5.x, ainsi que les composants qu'Amazon EMR installe avec Zeppelin.

Pour connaître la version des composants installés avec Zeppelin dans cette version, consultez Versions des composants de la version 5.36.2.

Informations sur la version de Zeppelin pour emr-5.36.2

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.36.2	Zeppelin 0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
		client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Rubriques

- Éléments à prendre en compte lors de l'utilisation de Zeppelin sur Amazon EMR
- Historique des versions de Zeppelin

Éléments à prendre en compte lors de l'utilisation de Zeppelin sur Amazon EMR

- Connectez-vous à Zeppelin à l'aide de la même méthode de tunnels SSH pour vous connecter aux autres serveurs web sur le nœud principal. Le serveur Zeppelin est disponible sur le port 8890.
- Zeppelin sur les versions Amazon EMR 5.0.0. et ultérieures prend en charge l'authentification Shiro.
- Zeppelin sur Amazon EMR version 5.8.0 et ultérieure prend en charge l'utilisation du catalogue de données AWS Glue comme métastore pour Spark SQL. Pour plus d'informations, consultez Utilisation du catalogue de données AWS Glue en tant que metastore pour Spark SQL.
- Zeppelin n'utilise pas certains des paramètres définis dans le fichier de configuration `spark-defaults.conf` de votre cluster (même s'il demande à YARN d'allouer dynamiquement des exécuteurs si vous avez défini `spark.dynamicAllocation.enabled` sur `true`. Vous devez

définir des paramètres de l'exécuteur (comme la mémoire et les noyaux) en cliquant sur l'onglet Interpreter (Interprète) de Zeppelin, puis redémarrer l'interprète pour en permettre l'utilisation.

- Les versions 6.10.0 et supérieures d'Amazon EMR prennent en charge l'intégration d'Apache Zeppelin avec Apache Flink. Pour plus d'informations, consultez Travailler avec les jobs Flink de Zeppelin dans Amazon EMR.
- Zeppelin sur Amazon EMR ne prend pas en charge l'interpréteur SparkR.

Historique des versions de Zeppelin

Le tableau suivant répertorie la version de Zeppelin incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

Informations sur la version de Zeppelin

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-7,9.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.8.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
		httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7,7.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-7.6.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.5.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-7.4.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.3.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-7.2.0	0.10.10.10.11.	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.36.2	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-7.1.0	0.10.10.10.11.	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.0.0	0.10.10.10.11.	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.15.0	0.10.10.10.11.	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.14.0	0.10.10.10.11.	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.13.0	0.10.10.10.11.	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.12.0	0.10.10.10.11.	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.11.1	0.10.10.10.11.	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.11.0	0.10.10.10.11.	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.10.1	0.10.10.10.11.	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.10.0	0.10.10.10.11.	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.9.1	0.10.10.10.11.	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.9.0	0.10.10.10.11.	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.8.1	0.10.10.10.11.	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.8.0	0.10.10.10.11.	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.7.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.36.1	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.36.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.6.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.35.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.5.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.4.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.3.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.3.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.2.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.2.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.1.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.1.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.0.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-6.0.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.34.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.33.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.33.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.32.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.32.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.31.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.31.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.30.2	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.30.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.30.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
mr-5.29.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.28.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.28.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.27.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.27.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.26.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.25.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.24.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.24.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.23.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.23.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.22.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.21.2	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.21.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.21.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.20.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.20.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.19.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.19.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.18.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.18.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.17.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.17.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.17.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.16.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.16.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.15.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.15.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.14.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.14.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.14.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.13.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.13.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.12.3	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.12.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.12.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.12.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.11.4	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.11.3	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.11.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.11.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.11.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.10.1	0.7.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.10.0	0.7.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.9.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.9.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.8.3	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.8.2	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.8.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.8.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.7.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.7.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.6.1	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.6.0	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.4	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.5.3	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.2	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.5.1	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.0	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.4.1	0.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.4.0	0.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.3.2	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.3.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.3.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.2.3	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.2.2	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.2.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.2.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.1.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.1.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.0.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.0.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.0.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Étiquette de version Amazon EMR	Version Zeppelin	Composants installés avec Zeppelin
emr-5.0.0	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Apache ZooKeeper

Apache ZooKeeper est un service centralisé permettant de gérer les informations de configuration, de nommer, de fournir une synchronisation distribuée et de fournir des services de groupe. Pour plus d'informations sur ZooKeeper, consultez http://zookeeper.apache.org/.

Le tableau suivant répertorie la version ZooKeeper incluse dans la dernière version d'Amazon EMR série 7.x, ainsi que les composants avec lesquels Amazon EMR installe. ZooKeeper

Pour connaître la version des composants installés avec ZooKeeper cette version, consultez les versions des composants de la version 7.9.0.

ZooKeeper informations de version pour emr-7.9.0

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-7.9.0	ZooKeeper 3.9.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Le tableau suivant répertorie la version ZooKeeper incluse dans la dernière version d'Amazon EMR série 6.x, ainsi que les composants avec lesquels Amazon EMR installe. ZooKeeper

Pour connaître la version des composants installés avec ZooKeeper cette version, consultez les versions des composants de la version 6.15.0.

ZooKeeper informations sur la version pour emr-6.15.0

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-6.15.0	ZooKeeper 3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Le tableau suivant répertorie la version ZooKeeper incluse dans la dernière version d'Amazon EMR série 5.x, ainsi que les composants avec lesquels Amazon EMR installe. ZooKeeper

Pour connaître la version des composants installés avec ZooKeeper cette version, consultez les versions des composants de la version 5.36.2.

ZooKeeper informations de version pour emr-5.36.2

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.36.2	ZooKeeper 3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Rubriques

- ZooKeeper historique des versions
- ZooKeeper notes de mise à jour par version

ZooKeeper historique des versions

Le tableau suivant répertorie la version ZooKeeper incluse dans chaque version d'Amazon EMR, ainsi que les composants installés avec l'application. Pour connaître les versions des composants de chaque version, consultez la section Version des composants de votre version dans Versions Amazon EMR 7.x, Versions Amazon EMR 6.x ou Versions Amazon EMR 5.x.

ZooKeeper informations sur la version

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-7.9.0	3.9.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.8.0	3.9.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-7,7.0	3.9.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.6.0	3.9.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.5.0	3.9.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-7.4.0	3.9.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.3.0	3.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.2.0	3.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.36.2	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.1.0	3.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.0.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-6.15.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.14.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.13.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-6.12.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.11.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.11.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-6.10.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.10.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.9.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-6.9.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.8.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.8.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-6.7.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.36.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.36.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-6.6.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.35.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.5.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-6.4.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.3.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.3.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-6.2.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.2.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.1.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-6.1.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.0.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.0.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.34.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.33.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.33.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.32.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.32.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.31.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.31.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.30.2	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.30.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.30.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
mr-5.29.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.28.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.28.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.27.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.27.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.26.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.25.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.24.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.24.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.23.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.23.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.22.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.21.2	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.21.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.21.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.20.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.20.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.19.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.19.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.18.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.18.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.17.2	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.17.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.17.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.16.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.16.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.15.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.15.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.14.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.14.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.14.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.13.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.13.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.12.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.4	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.11.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.11.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.10.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.10.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.9.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.9.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.8.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.7.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.7.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.6.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.6.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.5.4	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.5.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.4.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.4.0	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.3.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.3.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.3.0	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.2.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.1.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.1.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.3	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Étiquette de version Amazon EMR	ZooKeeper Version	Composants installés avec ZooKeeper
emr-5.0.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

ZooKeeper notes de mise à jour par version

Consultez les sections suivantes pour les notes de mise à jour complètes.

Amazon EMR 7.8.0 - notes de mise à jour ZooKeeper

Amazon EMR 7.8.0 - Modifications ZooKeeper

Mise à niveau — Zookeeper est mis à jour vers la version 3.9.3.

Amazon EMR 7.5.0 : notes de mise à jour ZooKeeper

Amazon EMR 7.5.0 - Modifications ZooKeeper

- À partir de EMR-7.5.0, Zookeeper est configuré sur Java 17 par défaut au moment de l'exécution. Pour utiliser une version différente pour le runtime Java, remplacez les paramètres JVM du `zookeeper-server` fichier et définissez `JAVA_HOME` la version souhaitée.

Amazon EMR 7.4.0 : notes de mise à jour ZooKeeper

Type	Description
Upgrade	Zookeeper est mise à jour vers la version 3.9.2.
Amélioration	ZOOKEEPER-4778 : corrigez jetty, netty et logback pour supprimer les vulnérabilités les plus graves.
Amélioration	ZOOKEEPER-4799 : Corrige le CVE-2024-23944.
Correctif de bogue.	Le support de ZOOKEEPER-4415 : TLSv1 .3 est ajouté.

Amazon EMR 7.4.0 - Fonctionnalités ZooKeeper

- Les clusters dotés d'une configuration de sécurité de chiffrement en transit fournie auront le protocole TLS activé par défaut pour les communications avec Quorum et le serveur d'administration Zookeeper. De plus, le port sécurisé TLS est exposé à 2281 pour la communication du client avec Zookeeper.

Connecteurs et utilitaires

Amazon EMR fournit un certain nombre de connecteurs et d'utilitaires permettant d'accéder à d'autres AWS services en tant que sources de données. Vous pouvez généralement accéder aux données de ces services au sein d'un programme. Par exemple, vous pouvez spécifier un flux Kinesis dans une requête Hive, un script Pig ou une MapReduce application, puis opérer sur ces données.

Rubriques

- Exportation, importation, interrogation et jointure de tables dans DynamoDB à l'aide d'Amazon EMR
- Kinesis
- S3 DistCp (s3-dist-cp)
- Nettoyage après l'échec de DistCp tâches S3

Exportation, importation, interrogation et jointure de tables dans DynamoDB à l'aide d'Amazon EMR

> **Note**
>
> Le connecteur Amazon EMR-DynamoDB est open source sur. GitHub Pour de plus amples informations, veuillez consulter https://github.com/awslabs/emr-dynamodb-connector.

DynamoDB est un service de base de données NoSQL entièrement géré offrant des performances rapides et prévisibles avec une scalabilité simple. Les développeurs peuvent créer une table de base de données et développer sans limite son trafic de demandes ou son stockage. DynamoDB répartit automatiquement les données et le trafic de la table sur un nombre de serveurs approprié afin de gérer le volume de requêtes spécifié par le client et la quantité de données stockées, tout en assurant la cohérence et la rapidité des performances. Avec Amazon EMR et Hive, vous pouvez traiter rapidement et efficacement de grandes quantités de données, telles que les données stockées dans DynamoDB. Pour plus d'informations sur DynamoDB, consultez le manuel du développeur Amazon DynamoDB.

Apache Hive est une couche logicielle que vous pouvez utiliser pour interroger, mapper ou réduire les clusters à l'aide d'un langage de requête simplifié de type SQL, appelé HiveQL. Il s'exécute sur l'architecture Hadoop. Pour plus d'informations sur Hive et HiveQL, consultez le Manuel du langage HiveQL. Pour plus d'informations sur Hive et Amazon EMR, consultez Apache Hive .

Vous pouvez utiliser Amazon EMR avec une version personnalisée de Hive qui inclut une connectivité à DynamoDB pour effectuer des opérations sur les données stockées dans DynamoDB :

- Chargement des données DynamoDB dans le système de fichier distribué Hadoop (HDFS) et son utilisation en tant qu'entrée dans un cluster Amazon EMR.
- Interrogation de données DynamoDB en direct à l'aide d'instructions de type SQL (HiveQL).
- Jointure de données stockées dans DynamoDB et exportation ou interrogation par rapport aux données jointes.
- Exportation de données stockées dans DynamoDB vers Amazon S3.
- Importation de données stockées dans Amazon S3 vers DynamoDB.

> **Note**
>
> Le connecteur Amazon EMR-DynamoDB ne prend pas en charge les clusters configurées pour utiliser l'authentification Kerberos.

Pour exécuter chacune des tâches suivantes, vous allez lancer un cluster Amazon EMR, spécifier l'emplacement des données dans DynamoDB et émettre des commandes Hive pour manipuler les données dans DynamoDB.

Il existe plusieurs manières de lancer un cluster Amazon EMR : vous pouvez utiliser la console Amazon EMR, l'interface de ligne de commande (CLI), ou vous pouvez programmer votre cluster à l'aide d'un AWS SDK ou de l'API Amazon EMR. Vous pouvez également choisir d'exécuter un cluster Hive de façon interactive ou à partir d'un script. Dans cette section, nous allons vous montrer comment lancer un cluster Hive interactif à partir de la console et de l'interface de ligne de commande Amazon EMR.

L'utilisation interactive de Hive est un excellent moyen de tester les performances des requêtes et de régler votre application. Une fois que vous avez défini un ensemble de commandes Hive qui s'exécuteront régulièrement, envisagez de créer un script Hive qu'Amazon EMR peut exécuter automatiquement.

> **Warning**
>
> Les opérations Amazon EMR de lecture et d'écriture sur une table DynamoDB desservent votre débit réservé établi, en augmentant potentiellement la fréquence des exceptions de débit réservé. Pour les demandes importantes, Amazon EMR implémente de nouvelles tentatives avec un backoff exponentiel pour gérer la charge des demandes sur la table DynamoDB. L'exécution simultanée de travaux Amazon EMR avec le reste du trafic peut entraîner un dépassement du niveau de débit réservé alloué. Vous pouvez contrôler cela en vérifiant la ThrottleRequestsmétrique sur Amazon CloudWatch. Si la charge de la demande est trop élevée, vous pouvez relancer le cluster et diminuer la valeur Paramètre de pourcentage de lecture ou Paramètre de pourcentage d'écriture afin de limiter les opérations Amazon EMR. Pour de plus amples informations sur les paramètres de débit DynamoDB, veuillez consulter Débit alloué.
>
> Si une table est configurée pour le mode à la demande, vous devez la remettre en mode provisionné avant d'exécuter une opération d'exportation ou d'importation. Les pipelines ont besoin d'un ratio de débit pour calculer les ressources à utiliser à partir d'une DynamoDBtable. Le mode à la demande supprime le débit provisionné. Pour provisionner la capacité de débit, vous pouvez utiliser les métriques Amazon CloudWatch Events pour évaluer le débit agrégé utilisé par une table.

Rubriques

- Configuration d'une table Hive pour exécuter des commandes Hive
- Exemples de commandes Hive pour l'exportation, l'importation et l'interrogation de données dans DynamoDB
- Optimisation des performances pour les opérations Amazon EMR dans DynamoDB

Configuration d'une table Hive pour exécuter des commandes Hive

Apache Hive est une application d'entrepôt des données que vous pouvez utiliser pour interroger les données contenues dans les clusters Amazon EMR à l'aide d'un langage de type SQL. Pour plus d'informations sur Hive, consultez le site http://hive.apache.org/.

La procédure suivante suppose que vous avez déjà créé un cluster et que vous avez spécifié une paire de EC2 clés Amazon. Pour savoir comment commencer à créer des clusters, consultez la section Premiers pas avec Amazon EMR dans le Guide de gestion Amazon EMR.

Configurer Hive pour utiliser MapReduce

Lorsque vous utilisez Hive sur Amazon EMR pour interroger les tables DynamoDB, des erreurs peuvent se produire si Hive utilise le moteur d'exécution par défaut, Tez. C'est pourquoi, lorsque vous créez un cluster avec Hive qui s'intègre à DynamoDB comme décrit dans cette section, nous vous recommandons d'utiliser une classification de configuration qui définit Hive à utiliser. MapReduce Pour de plus amples informations, veuillez consulter Configuration des applications.

L'extrait suivant montre la classification de configuration et les propriétés à utiliser pour définir MapReduce comme moteur d'exécution pour Hive :

```
[
                {
                    "Classification": "hive-site",
                    "Properties": {
                        "hive.execution.engine": "mr"
                    }
                }
            ]
```

Pour exécuter les commandes Hive de façon interactive

1. Connectez-vous au nœud maître. Pour plus d'informations, consultez Connexion au nœud principal à l'aide de SSH dans le Guide de gestion d'Amazon EMR.
2. A l'invite de commande pour le nœud maître actuel, tapez `hive`.

   Vous devez voir une invite Hive : `hive>`
3. Entrez une commande Hive qui mappe une table de l'application Hive aux données dans DynamoDB. Ce tableau agit comme référence pour les données stockées dans Amazon DynamoDB ; les données ne sont pas stockées localement dans Hive et les requêtes utilisant cette table s'exécutent sur les données en direct dans DynamoDB et consomment la capacité en lecture ou en écriture de la table chaque fois qu'une commande est exécutée. Si vous prévoyez d'exécuter plusieurs commandes Hive sur le même ensemble de données, pensez d'abord à l'exporter.

   L'exemple suivant montre la syntaxe pour mapper une table Hive à une table DynamoDB.

   ```
   CREATE EXTERNAL TABLE hive_tablename
    (hive_column1_name column1_datatype, hive_column2_name column2_datatype...)
   ```

```
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
"dynamodb.column.mapping" =
 "hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name..."
```

Lorsque vous créez une table dans Hive depuis DynamoDB, vous devez la créer en tant que table externe à l'aide du mot clé `EXTERNAL`. La différence entre les tables internes et externes est que les données des tables internes sont supprimées lorsqu'une table interne est supprimée. Il ne s'agit pas du comportement souhaité lorsque vous êtes connecté à Amazon DynamoDB, si bien que seules les tables externes sont prises en charge.

Par exemple, la commande Hive suivante crée une table nommée hivetable1 dans Hive qui fait référence à la table DynamoDB nommée dynamodbtable1. La table DynamoDB dynamodbtable1 possède un schéma de clé primaire. hash-and-range L'élément de clé de hachage est `name` (type de chaîne), l'élément de clé de plage est `year` (type numérique) et chaque élément a une valeur d'attribut pour `holidays` (type string set).

```
CREATE EXTERNAL TABLE hivetable1 (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");
```

La ligne 1 utilise l'instruction HiveQL `CREATE EXTERNAL TABLE`. Pour hivetable1, vous devez établir une colonne pour chaque paire nom-valeur d'attribut de la table DynamoDB et fournir le type de données. Ces valeurs ne sont pas sensibles à la casse et vous pouvez donner aux colonnes le nom de votre choix (à l'exception des mots réservés).

La ligne 2 utilise l'instruction `STORED BY`. La valeur de `STORED BY` est le nom de la classe qui gère la connexion entre Hive et DynamoDB. Cette valeur doit être définie sur `'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'`.

La ligne 3 utilise l'instruction `TBLPROPERTIES` pour associer « hivetable1 » à la table et au schéma appropriés dans DynamoDB. Fournissez à `TBLPROPERTIES` les valeurs des paramètres `dynamodb.table.name` et `dynamodb.column.mapping`. Ces valeurs sont sensibles à la casse.

> **Note**
>
> Tous les noms d'attribut DynamoDB de la table doivent avoir les colonnes correspondantes dans la table Hive. En fonction de votre version d'Amazon EMR, les scénarios suivants se produisent si le one-to-one mappage n'existe pas :
>
> - Sur Amazon EMR version 5.27.0 et versions ultérieures, le connecteur dispose de validations qui garantissent un mappage one-to-one entre les noms d'attributs DynamoDB et les colonnes de la table Hive. Une erreur se produira si le one-to-one mappage n'existe pas.
> - Sur Amazon EMR version 5.26.0 et antérieure, la table Hive ne contient pas de paire nom-valeur de DynamoDB. Si vous ne mappez pas les attributs de clé primaire DynamoDB, Hive génère une erreur. Si vous ne mappez pas un attribut de clé non primaire, aucune erreur n'est générée, mais vous ne voyez pas les données de la table Hive. Si les types de données ne correspondent pas, la valeur est null.

Ensuite, vous pouvez commencer à exécuter les opérations Hive sur hivetable1. Les requêtes exécutées sur hivetable1 sont exécutées en interne sur la table DynamoDB dynamodbtable1 de votre compte DynamoDB et consomment des unités de lecture ou d'écriture à chaque exécution.

Lorsque vous exécutez les requêtes Hive sur une table DynamoDB, vous devez vous assurer que vous avez alloué une quantité suffisante d'unités de capacité de lecture.

Par exemple, supposons que vous ayez provisionné 100 unités de capacité en lecture pour votre table DynamoDB. Vous pourrez ainsi effectuer 100 lectures, ou 409 600 octets, par seconde. Si cette table contient 20 Go de données (21 474 836 480 octets) et que votre requête Hive exécute une analyse complète de la table, vous pouvez estimer la durée d'exécution de la requête :

21 474 836 480 / 409 600 = 52 429 secondes = 14,56 heures

Le seul moyen de diminuer le temps nécessaire consiste à ajuster les unités de capacité de lecture sur la table DynamoDB source. L'ajout de nœuds Amazon EMR supplémentaires n'y changera rien.

Dans la sortie Hive, le pourcentage d'achèvement est mis à jour lorsqu'un ou plusieurs processus de mappeurs sont terminés. Pour une table DynamoDB volumineuse avec un paramètre de faible capacité allouée en lecture, la sortie du pourcentage d'achèvement peut ne pas être mise à jour pendant longtemps. Dans le cas décrit ci-dessus, la tâche apparaît complète à 0 % pendant plusieurs

heures. Pour un état plus détaillé de l'avancement de votre tâche, accédez à la console Amazon EMR ; vous serez en mesure d'afficher l'état des tâches de chaque mappeur et des statistiques de lecture des données. Vous pouvez également vous connecter à l'interface Hadoop sur le nœud maître et voir les statistiques Hadoop. Vous visualisez ainsi l'état des tâches de chaque carte et certaines statistiques de lecture des données. Pour plus d'informations, consultez les rubriques suivantes :

- Interfaces web hébergées sur le nœud principal
- Afficher les interfaces web Hadoop

Pour plus d'informations sur les exemples d'instructions HiveQL permettant d'exécuter des tâches telles que l'exportation ou l'importation de données depuis DynamoDB et la jointure de tables, consultez Exemples de commandes Hive pour l'exportation, l'importation et l'interrogation de données dans DynamoDB.

Pour annuler une demande Hive

Lorsque vous exécutez une requête Hive, la réponse initiale du serveur inclut la commande pour annuler la demande. Pour annuler la demande à tout moment du processus, utilisez la Kill Command (Commande kill) de la réponse du serveur.

1. Entrez `Ctrl+C` pour quitter le client de ligne de commande.
2. À la commande shell, saisissez la Commande kill à partir de la réponse initiale du serveur à votre demande.

   Vous pouvez également exécuter la commande suivante depuis la ligne de commande du nœud principal pour supprimer la tâche Hadoop. Elle contient l'identifiant de la tâche Hadoop et peut être récupérée depuis l'interface utilisateur de Hadoop. *`job-id`*

   ```
   hadoop job -kill job-id
   ```

Types de données pour Hive et DynamoDB

Le tableau suivant indique les types de données Hive disponibles, le type DynamoDB par défaut auquel ils correspondent et les autres types DynamoDB auxquels ils peuvent également mapper.

Type Hive	Type DynamoDB par défaut	Autre(s) type(s) de DynamoDB
chaîne	string (S)	
bigint ou double	number (N)	
binary	binary (B)	
boolean	boolean (BOOL)	
array	list (L)	number set (NS), string set (SS) ou binary set (BS)
map<string,string>	article	map (M)
map<string,?>	map (M)	
	null (NULL)	

Si vous voulez écrire vos données Hive sous la forme d'un autre type DynamoDB correspondant, ou si vos données DynamoDB contiennent des valeurs d'attribut d'un autre type DynamoDB, vous pouvez spécifier la colonne et le type DynamoDB avec le paramètre `dynamodb.type.mapping`. L'exemple suivant montre la syntaxe permettant de spécifier un mappage d'un autre type.

```
CREATE EXTERNAL TABLE hive_tablename (hive_column1_name column1_datatype,
 hive_column2_name column2_datatype...)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
"dynamodb.column.mapping" =
 "hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name...",
"dynamodb.type.mapping" = "hive_column1_name:dynamodb_attribute1_datatype");
```

Le paramètre de mappage de type est facultatif et ne doit être spécifié que pour les colonnes qui utilisent d'autres types.

Par exemple, la commande Hive suivante crée une table nommée `hivetable2` qui fait référence à la table DynamoDB `dynamodbtable2`. Elle est équivalent à `hivetable1`, sauf qu'elle mappe la colonne `col3` au type string set (SS).

```
CREATE EXTERNAL TABLE hivetable2 (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable2",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays",
"dynamodb.type.mapping" = "col3:SS");
```

Dans Hive, `hivetable1` et `hivetable2` sont identiques. Toutefois, lorsque les données de ces tables sont écrites dans leurs tables DynamoDB correspondantes, `dynamodbtable1` contient des listes, tandis que `dynamodbtable2` contient des ensembles de chaînes.

Si vous souhaitez écrire des valeurs `null` Hive en tant qu'attributs de type DynamoDB pour `null`, vous pouvez le faire avec le paramètre `dynamodb.null.serialization`. L'exemple suivant montre la syntaxe permettant de spécifier la sérialisation `null`.

```
CREATE EXTERNAL TABLE hive_tablename (hive_column1_name column1_datatype,
 hive_column2_name column2_datatype...)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
"dynamodb.column.mapping" =
 "hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name...",
"dynamodb.null.serialization" = "true");
```

Le paramètre de sérialisation null est facultatif ; il est défini sur `false` s'il n'est pas spécifié. Notez que les attributs `null` DynamoDB sont lus en tant que valeurs `null` dans Hive, quel que soit le paramètre. Les collections Hive avec des valeurs `null` peuvent être écrites dans DynamoDB uniquement si le paramètre de sérialisation null est spécifié comme `true`. Sinon, une erreur Hive se produit.

Le type bigint de Hive est identique au type long de Java, et le type double de Hive est le même que le type double de Java en termes de précision. Cela signifie que si vous avez des données numériques stockées dans DynamoDB qui ont une précision supérieure à celle des types de données Hive, l'utilisation de Hive pour exporter, importer ou référencer les données DynamoDB peut entraîner une perte de précision ou un échec de la requête Hive.

Les exportations de type binaire de DynamoDB vers Amazon Simple Storage Service (Amazon S3) ou HDFS sont stockées sous la forme d'une chaîne codée en Base64. Si vous importez des données

d'Amazon S3 ou HDFS dans le type binaire DynamoDB, elles doivent être encodées en tant que chaîne Base64.

Options Hive

Vous pouvez définir les options Hive suivantes pour gérer le transfert de données hors d'Amazon DynamoDB. Ces options sont conservées uniquement pour la session Hive en cours. Si vous fermez l'invite de commande Hive et la rouvrez ultérieurement sur le cluster, ces paramètres sont rétablis à leurs valeurs par défaut.

Options Hive	Description
`dynamodb.throughput.read.percent`	Définissez la vitesse des opérations de lecture pour maintenir votre débit DynamoDB dans la plage allouée pour votre table. La valeur est comprise entre `0.1` et `1.5`, valeurs incluses. La valeur de 0,5 est la vitesse de lecture par défaut, ce qui signifie que Hive essaie de consommer la moitié des lectures provisionnées dans le cadre des ressources de la table. L'augmentation de cette valeur au-dessus de 0,5 augmente le taux de demandes de lecture. Sa diminution sous 0,5 réduit le taux de demandes de lecture. Cette vitesse de lecture est approximative. La vitesse de lecture réelle dépend de facteurs comme l'existence d'une distribution uniforme des clés dans DynamoDB. Si vous trouvez que votre débit alloué est fréquemment dépassé par l'opération Hive, ou si le trafic de lecture en direct est trop limité, réduisez cette valeur sous `0.5`. Si vous avez suffisamment de capacité et voulez une opération Hive plus rapide, définissez cette valeur au-dessus de `0.5`. Vous pouvez également sursouscrire en la définissant sur 1,5 si vous pensez qu'il existe des opérations d'entrée/sortie inutilisées disponibles.

Options Hive	Description
`dynamodb.throughput.write.percent`	Définissez la vitesse des opérations d'écriture pour maintenir votre débit DynamoDB dans la plage allouée pour votre table. La valeur est comprise entre `0.1` et `1.5`, valeurs incluses. La valeur de 0,5 est la vitesse d'écriture par défaut, ce qui signifie que Hive essaie d'utiliser la moitié des écritures provisionnées dans le cadre des ressources de la table. L'augmentation de cette valeur au-dessus de 0,5 augmente le taux de demandes d'écriture. Sa diminution sous 0,5 réduit le taux de demandes d'écriture. Cette vitesse d'écriture est approximative. La vitesse d'écriture réelle dépend de facteurs comme l'existence d'une distribution uniforme des clés dans DynamoDB. Si vous trouvez que votre débit alloué est fréquemment dépassé par l'opération Hive ou si le trafic d'écriture en direct est trop limité, réduisez cette valeur sous `0.5`. Si vous avez suffisamment de capacité et voulez une opération Hive plus rapide, définissez cette valeur au-dessus de `0.5`. Vous pouvez également sursouscrire en la définissant jusqu'à 1,5 si vous pensez qu'il existe des opérations d'entrée/sortie inutilisées disponibles ou qu'il s'agit du chargement des données initiales sur la table et qu'il n'y pas encore de trafic en direct.
`dynamodb.endpoint`	Spécifiez le point de terminaison pour le service DynamoDB. Pour plus d'information sur les points de terminaison DynamoDB disponibles, consultez Régions et points de terminaison.

Options Hive	Description
`dynamodb.max.map.tasks`	Spécifiez le nombre maximal de tâches de mappage lors de la lecture des données à partir de DynamoDB. Cette valeur doit être supérieure ou égale à 1.
`dynamodb.retry.duration`	Spécifiez le nombre de minutes à utiliser comme délai d'expiration pour répéter les commandes Hive. Cette valeur doit être un entier supérieur ou égal à 0. Le délai d'expiration par défaut est de deux minutes.

Ces options sont définies à l'aide de la commande `SET`, comme illustré dans l'exemple suivant.

```
SET dynamodb.throughput.read.percent=1.0;

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

Exemples de commandes Hive pour l'exportation, l'importation et l'interrogation de données dans DynamoDB

Les exemples suivants utilisent les commandes Hive afin d'effectuer des opérations telles que l'exportation de données vers Amazon S3 ou HDFS, l'importation de données vers DynamoDB, la jointure de tables, l'interrogation de tables, et plus encore.

Les opérations sur une table Hive référencent les données stockées dans DynamoDB. Les commandes Hive sont soumises aux paramètres de débit alloué de la table DynamoDB et les données extraites incluent les données écrites dans la table DynamoDB au moment où la demande d'opération Hive est traitée par DynamoDB. Si le processus de récupération des données prend du temps, certaines données retournées par la commande Hive peuvent avoir été mises à jour dans DynamoDB depuis que la commande Hive a commencé.

Les commandes Hive `DROP TABLE` et `CREATE TABLE` agissent uniquement sur les tables locales de Hive et ne créent ni ne suppriment de tables dans DynamoDB. Si votre requête Hive fait référence à une table dans DynamoDB, cette table doit déjà exister avant que vous exécutiez la requête. Pour

plus d'informations sur la création et la suppression de tables dans DynamoDB, consultez Utilisation de tables dans DynamoDB dans le Guide du développeur Amazon DynamoDB.

> **Note**
>
> Lorsque vous mappez une table Hive à un emplacement dans Amazon S3, ne la mappez pas au chemin racine du compartiment, s3://amzn-s3-demo-bucket, car cela peut provoquer des erreurs lorsque Hive écrit les données dans Amazon S3. Mappez plutôt la table à un sous-chemin du bucket, s3://amzn-s3-demo-bucket/mypath.

Exportation de données à partir de DynamoDB

Vous pouvez utiliser Hive pour exporter des données à partir de DynamoDB.

Pour exporter une table DynamoDB vers un compartiment Amazon S3

- Créez une table Hive qui référence les données stockées dans DynamoDB. Ensuite, vous pouvez appeler la commande INSERT OVERWRITE pour écrire les données dans un répertoire externe. Dans l'exemple suivant, *s3://amzn-s3-demo-bucket/path/subpath/* il s'agit d'un chemin valide dans Amazon S3. Ajustez les colonnes et les types de données dans la commande CREATE pour faire correspondre les valeurs dans votre DynamoDB. Vous pouvez l'utiliser pour créer une archive de vos données DynamoDB dans Amazon S3.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


INSERT OVERWRITE DIRECTORY 's3://amzn-s3-demo-bucket/path/subpath/' SELECT *
FROM hiveTableName;
```

Pour exporter une table DynamoDB vers un compartiment Amazon S3 à l'aide de la compartimentation.

- Créez une table externe qui fait référence à un emplacement dans Amazon S3. Cela est présenté ci-dessous en tant que s3_export. Pendant l'appel CREATE, spécifiez le formatage des lignes de la table. Puis, lorsque vous utilisez INSERT OVERWRITE pour exporter les données de DynamoDB vers s3_export, les données sont écrites au format spécifié. Dans l'exemple suivant, les données sont écrites sous forme de valeurs séparées par des virgules (CSV).

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


CREATE EXTERNAL TABLE s3_export(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

Pour exporter une table DynamoDB vers un compartiment Amazon S3 sans spécifier de mappage de colonne

- Créez une table Hive qui référence les données stockées dans DynamoDB. Cet exemple est similaire au précédent, sauf que vous ne spécifiez pas un mappage de colonne. La table doit avoir exactement une colonne de type `map<string, string>`. Si vous créez ensuite une table `EXTERNAL` dans Amazon S3, vous pouvez appeler la commande `INSERT OVERWRITE` pour écrire les données de DynamoDB vers Amazon S3. Vous pouvez l'utiliser pour créer une archive de vos données DynamoDB dans Amazon S3. Comme il n'y a aucun mappage de colonnes, vous ne pouvez pas interroger les tables qui sont exportées de cette façon. L'exportation de données sans spécifier de mappage de colonne est disponible dans Hive 0.8.1.5 (ou version ultérieure), qui est prise en charge sur l'AMI Amazon EMR 2.2.x et les versions ultérieures.

```
CREATE EXTERNAL TABLE hiveTableName (item map<string,string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1");

CREATE EXTERNAL TABLE s3TableName (item map<string, string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE s3TableName SELECT *
FROM hiveTableName;
```

Pour exporter une table DynamoDB vers un compartiment Amazon S3 en utilisant la compartimentation des données

- Hive fournit plusieurs codecs de compression, que vous pouvez définir durant votre session Hive. Les données exportées sont ainsi compressées au format spécifié. L'exemple suivant compresse les fichiers exportés à l'aide de l'algorithme Lempel-Ziv-Oberhumer (LZO).

```
SET hive.exec.compress.output=true;
SET io.seqfile.compression.type=BLOCK;
SET mapred.output.compression.codec = com.hadoop.compression.lzo.LzopCodec;


CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


CREATE EXTERNAL TABLE lzo_compression_table (line STRING)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE lzo_compression_table SELECT *
FROM hiveTableName;
```

  Les codecs de compression disponibles sont les suivants :

- org.apache.hadoop.io.compress. GzipCodec
- org.apache.hadoop.io.compress. DefaultCodec
- com.hadoop.compression.lzo. LzoCodec
- com.hadoop.compression.lzo. LzopCodec
- org.apache.hadoop.io.compress. BZip2Codec
- org.apache.hadoop.io.compress. SnappyCodec

Pour exporter une table DynamoDB vers HDFS

- Utilisez la commande Hive suivante, où se *hdfs:///directoryName* trouve un chemin HDFS valide et *hiveTableName* une table dans Hive qui fait référence à DynamoDB. Cette opération d'exportation est plus rapide que l'exportation d'une table DynamoDB vers Amazon S3, car Hive 0.7.1.1 utilise HDFS comme étape intermédiaire lors de l'exportation des données vers Amazon S3. L'exemple suivant montre aussi comment définir `dynamodb.throughput.read.percent` sur 1.0 afin d'augmenter le taux de demande de lecture.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

SET dynamodb.throughput.read.percent=1.0;

INSERT OVERWRITE DIRECTORY 'hdfs:///directoryName' SELECT * FROM hiveTableName;
```

  Vous pouvez également exporter les données vers HDFS à l'aide de la mise en forme et de la compression comme indiqué ci-dessus pour l'exportation vers Amazon S3. Pour ce faire, remplacez simplement le répertoire Amazon S3 dans les exemples ci-dessus par un répertoire HDFS.

Pour lire les données de caractères UTF-8 non affichables dans Hive

- Vous pouvez lire et écrire des données de caractères UTF-8 non affichables avec Hive en utilisant la clause `STORED AS SEQUENCEFILE` lorsque vous créez la table. A SequenceFile est le format de fichier binaire Hadoop ; vous devez utiliser Hadoop pour lire ce fichier. L'exemple suivant montre comment exporter des données de DynamoDB vers Amazon S3. Vous pouvez utiliser cette fonctionnalité pour gérer les caractères codés UTF-8 non affichables.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


CREATE EXTERNAL TABLE s3_export(a_col string, b_col bigint, c_col array<string>)
STORED AS SEQUENCEFILE
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

Importation de données dans DynamoDB

Lorsque vous écrivez des données sur DynamoDB à l'aide de Hive, vous devez vous assurer que le nombre d'unités de capacité d'écriture est supérieur au nombre de mappeurs dans le cluster. Par exemple, les clusters qui s'exécutent sur des EC2 instances m1.xlarge produisent 8 mappeurs par instance. Dans le cas d'un cluster qui possède 10 instances, cela signifie un total de 80 mappeurs. Si vos unités de capacité en écriture ne sont pas supérieures au nombre de mappeurs dans le cluster, l'opération d'écriture Hive peut consommer la totalité du débit d'écriture ou tenter d'utiliser plus de débit qu'il n'en est alloué. Pour plus d'informations sur le nombre de mappeurs produits par chaque type d' EC2 instance, consultezConfiguration de Hadoop.

Le nombre de mappeurs dans Hadoop est contrôlé par les divisions en entrée. S'il y a trop peu de divisions, votre commande d'écriture peut ne pas être en mesure d'utiliser tout le débit d'écriture disponible.

Si un élément avec la même clé existe dans la table DynamoDB cible, il est remplacé. Si aucun élément avec la clé n'existe dans la table DynamoDB cible, l'élément est inséré.

Pour importer une table d'Amazon S3 vers DynamoDB

- Vous pouvez utiliser Amazon EMR (Amazon EMR) et Hive pour écrire des données d'Amazon S3 vers DynamoDB.

```
CREATE EXTERNAL TABLE s3_import(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

INSERT OVERWRITE TABLE hiveTableName SELECT * FROM s3_import;
```

Pour importer une table depuis un compartiment Amazon S3 vers DynamoDB sans spécifier de mappage de colonne

- Créez une table `EXTERNAL` qui référence les données stockées dans Amazon S3 et précédemment exportées depuis DynamoDB. Avant de les importer, vérifiez que la table existe dans DynamoDB et qu'elle possède le même schéma de clés que la table DynamoDB précédemment exportée. En outre, la table doit avoir une seule colonne de type `map<string, string>`. Si vous créez ensuite une table Hive liée à DynamoDB, vous pouvez appeler la commande `INSERT OVERWRITE` pour écrire les données depuis Amazon S3 vers DynamoDB. Comme il n'existe aucun mappage de colonne, vous ne pouvez pas interroger les tables qui sont importées de cette façon. L'importation de données sans spécifier un mappage de colonne est disponible dans Hive 0.8.1.5 ou version ultérieure, qui est pris en charge sur Amazon EMR AMI 2.2.3.x et version ultérieure.

```
CREATE EXTERNAL TABLE s3TableName (item map<string, string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

CREATE EXTERNAL TABLE hiveTableName (item map<string,string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1");
```

```
INSERT OVERWRITE TABLE hiveTableName SELECT *
FROM s3TableName;
```

Pour importer une table de HDFS vers DynamoDB

- Vous pouvez utiliser Amazon EMR et Hive pour écrire des données de HDFS vers DynamoDB.

```
CREATE EXTERNAL TABLE hdfs_import(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 'hdfs:///directoryName';

CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

INSERT OVERWRITE TABLE hiveTableName SELECT * FROM hdfs_import;
```

Interrogation de données dans DynamoDB

Les exemples suivants illustrent les différentes façons dont vous pouvez utiliser Amazon EMR pour interroger les données stockées dans DynamoDB.

Pour trouver la plus grande valeur pour une colonne mappée (**max**)

- Utilisez les commandes de Hive comme suit. Dans la première commande, l'instruction CREATE crée une table Hive qui référence les données stockées dans DynamoDB. L'instruction SELECT utilise ensuite cette table pour interroger les données stockées dans DynamoDB. L'exemple suivant recherche la plus grande commande passée par un client donné.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
```

```
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

SELECT max(total_cost) from hive_purchases where customerId = 717;
```

Pour regrouper les données à l'aide de la clause **GROUP BY**

- Vous pouvez utiliser la clause GROUP BY pour collecter les données de plusieurs enregistrements. Cette option est souvent utilisée avec une fonction d'agrégation comme somme, décompte, min ou max. L'exemple suivant retourne une liste des commandes les plus importantes des clients ayant passé plus de trois commandes.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

SELECT customerId, max(total_cost) from hive_purchases GROUP BY customerId HAVING
 count(*) > 3;
```

Pour joindre deux tables DynamoDB

- L'exemple suivant mappe deux tables Hive avec des données stockées dans DynamoDB. Il appelle ensuite une jointure entre ces deux tables. La jointure est calculée sur le cluster, puis renvoyée. La jointure n'a pas lieu dans DynamoDB. Cet exemple renvoie une liste de clients et de leurs achats pour ceux ayant passé plus de deux commandes.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");
```

```
CREATE EXTERNAL TABLE hive_customers(customerId bigint, customerName string,
 customerAddress array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Customers",
"dynamodb.column.mapping" =
 "customerId:CustomerId,customerName:Name,customerAddress:Address");

Select c.customerId, c.customerName, count(*) as count from hive_customers c
JOIN hive_purchases p ON c.customerId=p.customerId
GROUP BY c.customerId, c.customerName HAVING count > 2;
```

Pour joindre deux tables de différentes sources

- Dans l'exemple suivant, Customer_S3 est une table Hive qui charge un fichier CSV stocké dans Amazon S3 et hive_purchases est une table qui fait référence aux données de DynamoDB. L'exemple suivant joint les données client stockées comme fichier CSV dans Amazon S3 et les données relatives aux commandes stockées dans DynamoDB pour renvoyer un ensemble de données représentant les commandes passées par les clients ayant « Miller » dans leur nom.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

CREATE EXTERNAL TABLE Customer_S3(customerId bigint, customerName string,
 customerAddress array<String>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

Select c.customerId, c.customerName, c.customerAddress from
Customer_S3 c
JOIN hive_purchases p
ON c.customerid=p.customerid
where c.customerName like '%Miller%';
```

> **Note**
>
> Dans les exemples précédents, les instructions CREATE TABLE ont été incluses dans chaque exemple pour des raisons de clarté et d'exhaustivité. Lorsque vous exécutez plusieurs requêtes ou opérations d'exportation sur une table Hive donnée, vous ne devez créer la table qu'une seule fois, au début de la session Hive.

Optimisation des performances pour les opérations Amazon EMR dans DynamoDB

Les opérations Amazon EMR sur une table DynamoDB sont considérées comme des opérations de lecture, et sont soumises aux paramètres de débit approvisionnés de la table. Amazon EMR implémente sa propre logique pour l'équilibrage de la charge sur votre table DynamoDB, afin de minimiser la possibilité de dépassement du débit approvisionné. À la fin de chaque requête Hive, Amazon EMR renvoie des informations sur le cluster utilisé pour traiter la requête, y compris le nombre de fois où votre débit alloué a été dépassé. Vous pouvez utiliser ces informations, ainsi que CloudWatch les mesures relatives à votre débit DynamoDB, pour mieux gérer la charge sur votre table DynamoDB lors des demandes suivantes.

Les facteurs suivants influencent les performances des requêtes Hive lorsque vous travaillez avec des tables DynamoDB.

Unités de capacité de lecture allouées

Lorsque vous exécutez les requêtes Hive sur une table DynamoDB, vous devez vous assurer que vous avez alloué une quantité suffisante d'unités de capacité de lecture.

Par exemple, supposons que vous ayez provisionné 100 unités de capacité en lecture pour votre table DynamoDB. Vous pourrez ainsi effectuer 100 lectures, ou 409 600 octets, par seconde. Si cette table contient 20 Go de données (21 474 836 480 octets) et que votre requête Hive exécute une analyse complète de la table, vous pouvez estimer la durée d'exécution de la requête :

21 474 836 480 / 409 600 = 52 429 secondes = 14,56 heures

Le seul moyen de diminuer le temps nécessaire consiste à ajuster les unités de capacité de lecture sur la table DynamoDB source. L'ajout de nœuds supplémentaires au cluster Amazon EMR ne sera d'aucune utilité.

Dans la sortie Hive, le pourcentage d'achèvement est mis à jour lorsqu'un ou plusieurs processus de mappeurs sont terminés. Pour une table DynamoDB volumineuse avec un paramètre de faible capacité allouée en lecture, la sortie du pourcentage d'achèvement peut ne pas être mise à jour pendant longtemps. Dans le cas décrit ci-dessus, la tâche apparaît complète à 0 % pendant plusieurs heures. Pour un état plus détaillé de l'avancement de votre tâche, accédez à la console Amazon EMR ; vous serez en mesure d'afficher l'état des tâches de chaque mappeur et des statistiques de lecture des données.

Vous pouvez également vous connecter à l'interface Hadoop sur le nœud maître et voir les statistiques Hadoop. Vous visualisez ainsi l'état des tâches de chaque carte et certaines statistiques de lecture des données. Pour plus d'informations, consultez Interfaces web hébergées sur le nœud principal dans le Guide de gestion d'Amazon EMR.

Paramètre de pourcentage de lecture

Par défaut, Amazon EMR gère la charge de requête sur votre table DynamoDB selon votre débit alloué actuel. Cependant, quand Amazon EMR renvoie des informations sur votre tâche avec un nombre élevé de réponses de dépassement du débit alloué, vous pouvez ajuster la vitesse de lecture par défaut à l'aide du paramètre `dynamodb.throughput.read.percent` lorsque vous configurez la table Hive. Pour plus d'informations sur la configuration du paramètre de pourcentage de lecture, consultez Options Hive.

Paramètre de pourcentage d'écriture

Par défaut, Amazon EMR gère la charge de requête sur votre table DynamoDB selon votre débit alloué actuel. Cependant, quand Amazon EMR renvoie des informations sur votre tâche avec un nombre élevé de réponses de dépassement du débit alloué, vous pouvez ajuster la vitesse d'écriture par défaut à l'aide du paramètre `dynamodb.throughput.write.percent` lorsque vous configurez la table Hive. Pour plus d'informations sur la configuration du paramètre de pourcentage d'écriture, consultez Options Hive.

Paramètre de durée de nouvelle tentative

Par défaut, Amazon EMR exécute à nouveau une requête Hive si elle n'a pas renvoyé de résultat dans les deux minutes, l'intervalle de nouvelle tentative par défaut. Vous pouvez ajuster cet intervalle en définissant le paramètre `dynamodb.retry.duration` lorsque vous exécutez une requête Hive. Pour plus d'informations sur la configuration du paramètre de pourcentage d'écriture, consultez Options Hive.

Nombre de tâches de mappage

Les démons de mappeur que Hadoop lance pour traiter vos requêtes afin d'exporter et d'interroger des données stockées dans DynamoDB sont limités à une vitesse de lecture maximale de 1 Mio par seconde afin de limiter la capacité de lecture utilisée. Si vous avez un débit alloué supplémentaire disponible sur DynamoDB, vous pouvez améliorer les performances de l'exportation Hive et des opérations d'interrogation en augmentant le nombre de démons de mappeur. Pour ce faire, vous pouvez soit augmenter le nombre d' EC2 instances de votre cluster, soit augmenter le nombre de démons de mappage exécutés sur chaque instance. EC2

Vous pouvez augmenter le nombre d' EC2 instances d'un cluster en arrêtant le cluster actuel et en le relançant avec un plus grand nombre d' EC2 instances. Vous spécifiez le nombre d' EC2 instances dans la boîte de dialogue Configurer les EC2 instances si vous lancez le cluster depuis la console Amazon EMR, ou avec l'`--num-instances`option si vous lancez le cluster depuis la CLI.

Le nombre de tâches cartographiques exécutées sur une instance dépend du type d' EC2 instance. Pour plus d'informations sur les types d' EC2 instances pris en charge et le nombre de mappeurs fournis par chacun d'entre eux, consultezConfiguration de la tâche. Là, vous trouverez une section « Configuration des tâches » pour chacune des configurations prises en charge.

Un autre moyen d'augmenter le nombre de démons de mappeur consiste à modifier le paramètre de configuration `mapreduce.tasktracker.map.tasks.maximum` de Hadoop sur une valeur plus élevée. Cela a l'avantage de vous donner plus de mappeurs sans augmenter le nombre ou la taille des EC2 instances, ce qui vous permet d'économiser de l'argent. L'inconvénient est que si cette valeur est trop élevée, les EC2 instances de votre cluster risquent de manquer de mémoire. Pour définir `mapreduce.tasktracker.map.tasks.maximum`, lancez le cluster et spécifiez une valeur pour `mapreduce.tasktracker.map.tasks.maximum` en tant que propriété de la classification de configuration mapred-site. Voici un exemple : Pour de plus amples informations, veuillez consulter Configuration des applications.

```
{
    "configurations": [
    {
        "classification": "mapred-site",
        "properties": {
            "mapred.tasktracker.map.tasks.maximum": "10"
        }
    }
    ]
```

```
}
```

Demandes de données en parallèle

Plusieurs demandes de données, à partir de plusieurs utilisateurs ou de plusieurs applications vers une seule table, peuvent diminuer la vitesse de lecture allouée et ralentir les performances.

Durée du processus

La cohérence des données dans DynamoDB dépend de l'ordre des opérations de lecture et d'écriture sur chaque nœud. Quand une requête Hive est en cours, une autre application peut charger de nouvelles données dans la table DynamoDB, voire modifier ou supprimer des données existantes. Dans ce cas, les résultats de la requête Hive peuvent ne pas tenir compte des modifications apportées aux données pendant l'exécution de la requête.

Eviter le dépassement de débit

Lors de l'exécution de requêtes Hive sur DynamoDB, veillez à ne pas dépasser votre débit alloué, vous risquez sinon de réduire la capacité nécessaire des appels à `DynamoDB::Get` de votre application. Pour éviter que cela ne se produise, vous devez surveiller régulièrement le volume de lecture et la limitation des appels aux applications `DynamoDB::Get` en consultant les journaux et les métriques de surveillance sur Amazon. CloudWatch

Durée de la demande

La planification des requêtes Hive qui accèdent à une table DynamoDB à un moment où la demande sur celle-ci est inférieure a pour effet d'améliorer les performances. Par exemple, si la plupart des utilisateurs de votre application vivent à San Francisco, vous pouvez choisir d'exporter les données chaque jour à 4h00 HNP, lorsque la majorité des utilisateurs dort, sans mettre à jour les enregistrements de votre base de données DynamoDB.

Tables basées sur le temps

Si les données sont organisées en tant que série de tables DynamoDB basées sur le temps, telles qu'une table par jour, vous pouvez exporter les données lorsque la table n'est plus active. Vous pouvez utiliser cette technique pour sauvegarder les données sur Amazon S3 en continu.

Données archivées

Si vous prévoyez d'exécuter plusieurs requêtes Hive sur les données stockées dans DynamoDB et que votre application peut tolérer les données archivées, il se peut que vous souhaitiez exporter les

données vers HDFS ou Amazon S3, et exécuter les requêtes Hive sur une copie des données au lieu de DynamoDB. Cela permet de conserver vos opérations de lecture et votre débit alloué.

Kinesis

Les clusters Amazon EMR peuvent lire et traiter directement les flux Amazon Kinesis à l'aide d'outils courants de l'écosystème Hadoop tels que Hive, Pig MapReduce, l'API de streaming Hadoop et Cascading. Vous pouvez également joindre les données en temps réel d'Amazon Kinesis aux données existantes sur Amazon S3, Amazon DynamoDB et HDFS dans un cluster en cours d'exécution. Vous pouvez charger directement les données d'Amazon EMR sur Amazon S3 ou DynamoDB pour les activités de post-traitement. Pour plus d’informations sur les caractéristiques principales du service Amazon Kinesis et ses tarifs, voir la page Amazon Kinesis.

Que puis-je faire avec l'intégration d'Amazon EMR et Amazon Kinesis ?

L'intégration entre Amazon EMR et Amazon Kinesis facilite considérablement certains scénarios, par exemple :

- Analyse de journaux de streaming – Vous pouvez analyser les journaux web de vos flux de diffusion continus afin de générer la liste des 10 principaux types d'erreurs par région, par navigateur et par domaine d'accès à intervalles de quelques minutes.
- Engagement client – Vous pouvez écrire des requêtes associant des données de parcours provenant d'Amazon Kinesis à des informations sur une campagne publicitaire stockées dans une table DynamoDB, dans le but d'identifier les catégories de publicité les plus efficaces parmi celles affichées sur des sites web donnés.
- Requêtes interactives ad-hoc – Vous pouvez charger périodiquement des données provenant de flux Amazon Kinesis dans HDFS et les mettre à disposition sous forme d'une table Impala locale permettant des requêtes analytiques, rapides et interactives.

Analyse de points de contrôle des flux Amazon Kinesis

Les utilisateurs peuvent exécuter des analyses régulières par lots de flux Amazon Kinesis dans ce que l'on appelle des itérations. Étant donné que les enregistrements de données de flux Amazon Kinesis sont récupérés à l'aide d'un numéro de séquence, des limites d'itérations sont définies par des numéros de début et de fin de séquences qu'Amazon EMR stocke dans une table DynamoDB. Par exemple, quand `iteration0` se termine, elle stocke le numéro de fin de séquence dans

DynamoDB afin que lorsque la tâche `iteration1` commence, elle puisse extraire les données suivantes du flux. Ce mappage d'itérations de données de flux s'appelle un point de contrôle. Pour plus d'informations, consultez Connecteur Kinesis.

Si une itération a été vérifiée et que la tâche a échoué lors du traitement d'une itération, Amazon EMR tente de retraiter les enregistrements de cette itération.

Les points de contrôle sont une fonctionnalité qui vous permet de :

- Commencer le traitement de données après un numéro de séquence traité par une requête précédente qui s'est exécutée sur le même flux de données et nom logique
- Retraiter le même lot de données à partir de Kinesis qui a été traité par une requête précédente

Pour activer les points de contrôle, définissez le paramètre `kinesis.checkpoint.enabled` sur `true` dans vos scripts. En outre, configurez les paramètres suivants :

Paramètre de configuration	Description
kinesis.checkpoint.metastore.table.name	Nom de la table DynamoDB où seront stockées les informations de point de contrôle
kinesis.checkpoint.metastore.hash.key.name	Nom de la clé de hachage pour la table DynamoDB
kinesis.checkpoint.metastore.hash.range.name	Nom de la clé de plage pour la table DynamoDB
kinesis.checkpoint.logical.name	Un nom logique pour le traitement actuel
kinesis.checkpoint.iteration.no	Numéro de l'itération pour le traitement associé au nom logique
kinesis.rerun.iteration.without.wait	Valeur booléenne qui indique si une itération ayant échoué peut être exécutée à nouveau sans attente l'expiration ; la valeur par défaut est `false`

Recommandations d'IOPS provisionnés pour les tables Amazon DynamoDB

Le connecteur Amazon EMR pour Amazon Kinesis utilise la base de données DynamoDB comme support pour les métadonnées de pointage. Vous devez créer une table dans DynamoDB avant de consommer des données dans un flux Amazon Kinesis avec un cluster Amazon EMR à intervalles contrôlés. La table doit se trouver dans la même région que votre cluster Amazon EMR. Voici des recommandations générales pour le nombre d'IOPS que vous devez mettre en service pour vos tables DynamoDB. `j` est le nombre maximum de tâches Hadoop (avec une combinaison de nombre d'itérations+nom logique différente) pouvant s'exécuter simultanément et `s` est le nombre maximum de partitions que toute tâche traitera :

Pour Read Capacity Units (Unités de capacité en lecture) : `j*s/5`

Pour Write Capacity Units (Unités de capacité en écriture) : `j*s`

Considérations sur les performances

Le débit de la partition Amazon Kinesis est directement proportionnel à la taille de l'instance des nœuds dans les clusters Amazon EMR et à la taille de l'enregistrement dans le flux. Nous vous recommandons d'utiliser des instances m5.xlarge ou plus importantes sur les nœuds maîtres et principaux.

Planifiez une analyse Amazon Kinesis avec Amazon EMR

Lorsque vous analysez des données sur un flux Amazon Kinesis actif, limité par des délais d'attente et une durée maximale pour l'itération, il est important que vous exécutiez l'analyse fréquemment pour collecter régulièrement des détails à partir du flux. Il existe plusieurs façons d'exécuter ces scripts et requêtes à intervalles réguliers ; nous vous conseillons d'utiliser AWS Data Pipeline pour les tâches répétitives de ce type. Pour plus d'informations, consultez AWS Data Pipeline PigActivityet consultez AWS Data Pipeline HiveActivityle Guide du AWS Data Pipeline développeur.

Migration du connecteur Spark Kinesis vers le SDK 2.x pour Amazon EMR 7.0

Le AWS SDK fournit un ensemble complet de bibliothèques permettant d'interagir avec APIs les services de AWS cloud computing, tels que la gestion des informations d'identification, la connexion aux services S3 et Kinesis. Le connecteur Spark Kinesis est utilisé pour consommer les données des flux de données Kinesis, et les données reçues sont transformées et traitées dans le moteur

d'exécution de Spark. Actuellement, ce connecteur est basé sur les versions 1.x du AWS SDK et Kinesis-client-library (KCL).

Dans le cadre de la migration vers le AWS SDK 2.x, le connecteur Spark Kinesis est également mis à jour en conséquence pour fonctionner avec le SDK 2.x. Dans la version 7.0 d'Amazon EMR, Spark contient la mise à niveau vers le SDK 2.x, qui n'est pas encore disponible dans la version communautaire d'Apache Spark. Si vous utilisez le connecteur Spark Kinesis sur une version antérieure à la version 7.0, migrez les codes de vos applications pour qu'ils s'exécutent sur le SDK 2.x avant de migrer vers Amazon EMR 7.0.

Guides de migration

Cette section décrit les étapes de migration d'une application vers la version mise à niveau du connecteur Spark Kinesis. Il inclut des guides de migration vers la bibliothèque client Kinesis (KCL) 2.x, des fournisseurs AWS d'informations d'identification et des clients de AWS service dans le SDK 2.x. AWS À titre de référence, il inclut également un exemple de WordCountprogramme utilisant le connecteur Kinesis.

Rubriques

- Migration de KCL de la version 1.x vers la version 2.x
- Migration des fournisseurs AWS d'informations d'identification du AWS SDK 1.x vers le SDK 2.x
- Migration des clients AWS de service du AWS SDK 1.x vers le SDK 2.x
- Exemples de code pour les applications de streaming
- Considérations relatives à l'utilisation de la version mise à niveau du connecteur Spark Kinesis

Migration de KCL de la version 1.x vers la version 2.x

- Niveau et dimensions des métriques dans **`KinesisInputDStream`**

  Lorsque vous instanciez un `KinesisInputDStream`, vous pouvez contrôler le niveau et les dimensions des métriques du flux. L'exemple suivant montre comment personnaliser ces paramètres avec KCL 1.x :

```
import
 com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibConfiguration
import com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel

val kinesisStream = KinesisInputDStream.builder
```

```
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(kinesisCheckpointInterval)
  .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
  .metricsLevel(MetricsLevel.DETAILED)

 .metricsEnabledDimensions(KinesisClientLibConfiguration.DEFAULT_METRICS_ENABLED_DIMENSIONS.a
  .build()
```

Dans KCL 2.x, ces paramètres de configuration ont des noms de package différents. Pour migrer vers la version 2.x :

1. Remplacez les instructions d'importation pour `com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibCon` et `com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel` par `software.amazon.kinesis.metrics.MetricsLevel` et `software.amazon.kinesis.metrics.MetricsUtil` respectivement.

   ```
   // import com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel
   import software.amazon.kinesis.metrics.MetricsLevel

   // import
    com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibConfiguration
   import software.amazon.kinesis.metrics.MetricsUtil
   ```

2. Remplacez la ligne `metricsEnabledDimensionsKinesisClientLibConfiguration.DEFAULT_METRICS_ENABL` par `metricsEnabledDimensionsSet(MetricsUtil.OPERATION_DIMENSION_NAME, MetricsUtil.SHARD_ID_DIMENSION_NAME)`.

Vous trouverez ci-dessous une version mise à jour du `KinesisInputDStream` avec un niveau de métrique et des dimensions de métriques personnalisés :

```
import software.amazon.kinesis.metrics.MetricsLevel
import software.amazon.kinesis.metrics.MetricsUtil

```

```
val kinesisStream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(kinesisCheckpointInterval)
  .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
  .metricsLevel(MetricsLevel.DETAILED)
  .metricsEnabledDimensions(Set(MetricsUtil.OPERATION_DIMENSION_NAME,
 MetricsUtil.SHARD_ID_DIMENSION_NAME))
  .build()
```

- Fonction de gestionnaire de messages dans `KinesisInputDStream`

  Lors de l'instanciation d'un `KinesisInputDStream`, vous pouvez également spécifier une « fonction de gestionnaire de messages » qui prend un enregistrement Kinesis et renvoie un objet générique T, au cas où vous souhaiteriez utiliser d'autres données incluses dans un enregistrement, comme une clé de partition.

  Dans KCL 1.x, la signature de la fonction de gestionnaire de messages est `Record => T`, où Record correspond à `com.amazonaws.services.kinesis.model.Record`. Dans KCL 2.x, la signature du gestionnaire est remplacée par :`KinesisClientRecord => T`, where is. KinesisClientRecord `software.amazon.kinesis.retrieval.KinesisClientRecord`

  Voici un exemple de spécification de gestionnaire de messages dans KCL 1.x :

```
import com.amazonaws.services.kinesis.model.Record


def addFive(r: Record): Int = JavaUtils.bytesToString(r.getData).toInt + 5
val stream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(Seconds(10))
  .storageLevel(StorageLevel.MEMORY_ONLY)
  .buildWithMessageHandler(addFive)
```

Pour migrer le gestionnaire de messages :

1. Remplacez l'instruction d'importation `com.amazonaws.services.kinesis.model.Record` par `software.amazon.kinesis.retrieval.KinesisClientRecord`.

```
// import com.amazonaws.services.kinesis.model.Record
import software.amazon.kinesis.retrieval.KinesisClientRecord
```

2. Mettez à jour la signature de la méthode du gestionnaire de messages.

```
//def addFive(r: Record): Int = JavaUtils.bytesToString(r.getData).toInt + 5
def addFive = (r: KinesisClientRecord) => JavaUtils.bytesToString(r.data()).toInt
 + 5
```

Voici un exemple mis à jour de spécification de gestionnaire de messages dans KCL 2.x :

```
import software.amazon.kinesis.retrieval.KinesisClientRecord


def addFive = (r: KinesisClientRecord) => JavaUtils.bytesToString(r.data()).toInt + 5
val stream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(Seconds(10))
  .storageLevel(StorageLevel.MEMORY_ONLY)
  .buildWithMessageHandler(addFive)
```

Pour plus d'informations sur la migration de KCL 1.x vers KCL 2.x, voir la rubrique Migration des consommateurs de KCL 1.x vers KCL 2.x.

Migration des fournisseurs AWS d'informations d'identification du AWS SDK 1.x vers le SDK 2.x

Les fournisseurs d'informations d'identification sont utilisés pour obtenir des AWS informations d'identification pour les interactions avec AWS. Il existe plusieurs modifications d'interface

et de classe liées aux fournisseurs d'informations d'identification dans le SDK 2.x. Ces modifications peuvent être consultées ici. Le connecteur Spark Kinesis a défini une interface (`org.apache.spark.streaming.kinesis.SparkAWSCredentials`) et des classes d'implémentation qui renvoient la version 1.x des fournisseurs d'informations d' AWS identification. Ces fournisseurs d'informations d'identification sont nécessaires lors de l'initialisation des clients Kinesis. Par exemple, si vous utilisez cette méthode `SparkAWSCredentials.provider` dans les applications, vous devrez mettre à jour les codes pour utiliser la version 2.x des fournisseurs AWS d'informations d'identification.

Voici un exemple d'utilisation des fournisseurs d'informations d'identification dans le AWS SDK 1.x :

```
import org.apache.spark.streaming.kinesis.SparkAWSCredentials
import com.amazonaws.auth.AWSCredentialsProvider

val basicSparkCredentials = SparkAWSCredentials.builder
    .basicCredentials("accessKey", "secretKey")
    .build()

val credentialProvider = basicSparkCredentials.provider
assert(credentialProvider.isInstanceOf[AWSCredentialsProvider], "Type should be
 AWSCredentialsProvider")
```

Pour migrer vers la version 2.x du SDK :

1. Remplacez l'instruction d'importation `com.amazonaws.auth.AWSCredentialsProvider` par `software.amazon.awssdk.auth.credentials.AwsCredentialsProvider`.

   ```
   //import com.amazonaws.auth.AWSCredentialsProvider
   import software.amazon.awssdk.auth.credentials.AwsCredentialsProvider
   ```

2. Mettez à jour les autres codes qui utilisent cette classe.

   ```
   import org.apache.spark.streaming.kinesis.SparkAWSCredentials
   import software.amazon.awssdk.auth.credentials.AwsCredentialsProvider

   val basicSparkCredentials = SparkAWSCredentials.builder
       .basicCredentials("accessKey", "secretKey")
       .build()

   val credentialProvider = basicSparkCredentials.provider
   ```

```
assert (credentialProvider.isInstanceOf[AwsCredentialsProvider], "Type should be
 AwsCredentialsProvider")
```

Migration des clients AWS de service du AWS SDK 1.x vers le SDK 2.x

AWS les clients de service ont des noms de package différents dans la version 2.x (c'est-à-dire`software.amazon.awssdk`), alors que le SDK 1.x les utilise. `com.amazonaws` Pour plus d'informations sur les modifications apportées aux clients, voir cette page. Si vous utilisez ces clients de service dans les codes, vous devez migrer les clients en conséquence.

Voici un exemple de création d'un client dans la version 1.x du SDK :

```
import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient
import com.amazonaws.services.dynamodbv2.document.DynamoDB

AmazonDynamoDB ddbClient = AmazonDynamoDBClientBuilder.defaultClient();
AmazonDynamoDBClient ddbClient = new AmazonDynamoDBClient();
```

Pour migrer vers la version 2.x :

1. Modifiez les instructions d'importation pour les clients de service. Prenons l'exemple des clients DynamoDB. Remplacez `com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient` ou `com.amazonaws.services.dynamodbv2.document.DynamoDB` par `software.amazon.awssdk.services.dynamodb.DynamoDbClient`.

   ```
   // import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient
   // import com.amazonaws.services.dynamodbv2.document.DynamoDB
   import software.amazon.awssdk.services.dynamodb.DynamoDbClient
   ```

2. Mettez à jour les codes qui initialisent les clients.

   ```
   // AmazonDynamoDB ddbClient = AmazonDynamoDBClientBuilder.defaultClient();
   // AmazonDynamoDBClient ddbClient = new AmazonDynamoDBClient();

   DynamoDbClient ddbClient = DynamoDbClient.create();
   DynamoDbClient ddbClient = DynamoDbClient.builder().build();
   ```

   Pour plus d'informations sur la migration du AWS SDK de la version 1.x vers la version 2.x, consultez la section Quelles sont les différences entre le SDK pour AWS Java 1.x et 2.x ?

Exemples de code pour les applications de streaming

```
import java.net.URI
import software.amazon.awssdk.auth.credentials.DefaultCredentialsProvider
import software.amazon.awssdk.http.apache.ApacheHttpClient
import software.amazon.awssdk.services.kinesis.KinesisClient
import software.amazon.awssdk.services.kinesis.model.DescribeStreamRequest
import software.amazon.awssdk.regions.Region
import software.amazon.kinesis.metrics.{MetricsLevel, MetricsUtil}

import org.apache.spark.SparkConf
import org.apache.spark.storage.StorageLevel
import org.apache.spark.streaming.{Milliseconds, StreamingContext}
import org.apache.spark.streaming.dstream.DStream.toPairDStreamFunctions
import org.apache.spark.streaming.kinesis.KinesisInitialPositions.Latest
import org.apache.spark.streaming.kinesis.KinesisInputDStream


object KinesisWordCountASLSDKV2 {

  def main(args: Array[String]): Unit = {
    val appName = "demo-app"
    val streamName = "demo-kinesis-test"
    val endpointUrl = "https://kinesis.us-west-2.amazonaws.com"
    val regionName = "us-west-2"

    // Determine the number of shards from the stream using the low-level Kinesis
 Client
    // from the AWS Java SDK.
    val credentialsProvider = DefaultCredentialsProvider.create
    require(credentialsProvider.resolveCredentials() != null,
      "No AWS credentials found. Please specify credentials using one of the methods
 specified " +
        "in https://docs.aws.amazon.com/sdk-for-java/latest/developer-guide/
credentials.html")
    val kinesisClient = KinesisClient.builder()
      .credentialsProvider(credentialsProvider)
      .region(Region.US_WEST_2)
      .endpointOverride(URI.create(endpointUrl))
      .httpClientBuilder(ApacheHttpClient.builder())
      .build()
    val describeStreamRequest = DescribeStreamRequest.builder()
      .streamName(streamName)
      .build()
```

```
    val numShards = kinesisClient.describeStream(describeStreamRequest)
      .streamDescription
      .shards
      .size


    // In this example, we are going to create 1 Kinesis Receiver/input DStream for
each shard.
    // This is not a necessity; if there are less receivers/DStreams than the number of
shards,
    // then the shards will be automatically distributed among the receivers and each
receiver
    // will receive data from multiple shards.
    val numStreams = numShards

    // Spark Streaming batch interval
    val batchInterval = Milliseconds(2000)

    // Kinesis checkpoint interval is the interval at which the DynamoDB is updated
with information
    // on sequence number of records that have been received. Same as batchInterval for
this
    // example.
    val kinesisCheckpointInterval = batchInterval

    // Setup the SparkConfig and StreamingContext
    val sparkConfig = new SparkConf().setAppName("KinesisWordCountASLSDKV2")
    val ssc = new StreamingContext(sparkConfig, batchInterval)

    // Create the Kinesis DStreams
    val kinesisStreams = (0 until numStreams).map { i =>
      KinesisInputDStream.builder
        .streamingContext(ssc)
        .streamName(streamName)
        .endpointUrl(endpointUrl)
        .regionName(regionName)
        .initialPosition(new Latest())
        .checkpointAppName(appName)
        .checkpointInterval(kinesisCheckpointInterval)
        .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
        .metricsLevel(MetricsLevel.DETAILED)
        .metricsEnabledDimensions(Set(MetricsUtil.OPERATION_DIMENSION_NAME,
MetricsUtil.SHARD_ID_DIMENSION_NAME))
        .build()
```

```
    }

    // Union all the streams
    val unionStreams = ssc.union(kinesisStreams)

    // Convert each line of Array[Byte] to String, and split into words
    val words = unionStreams.flatMap(byteArray => new String(byteArray).split(" "))

    // Map each word to a (word, 1) tuple so we can reduce by key to count the words
    val wordCounts = words.map(word => (word, 1)).reduceByKey(_ + _)

    // Print the first 10 wordCounts
    wordCounts.print()

    // Start the streaming context and await termination
    ssc.start()
    ssc.awaitTermination()
  }
}
```

Considérations relatives à l'utilisation de la version mise à niveau du connecteur Spark Kinesis

- Si vos applications utilisent la `Kinesis-producer-library` avec une version du JDK antérieure à la version 11, des exceptions peuvent se produire, comme `java.lang.NoClassDefFoundError: javax/xml/bind/DatatypeConverter`. Cela est dû au fait que EMR 7.0 est fourni avec le JDK 17 par défaut et que les modules J2EE ont été supprimés des bibliothèques standard depuis Java 11. Il est possible de résoudre le problème en ajoutant la dépendance suivante dans le fichier pom. Remplacez la version de la bibliothèque par une autre selon vos besoins.

  ```
  <dependency>
        <groupId>javax.xml.bind</groupId>
        <artifactId>jaxb-api</artifactId>
        <version>${jaxb-api.version}</version>
      </dependency>
  ```

- Le fichier jar du connecteur Spark Kinesis se trouve à l'emplacement suivant après la création d'un cluster EMR : `/usr/lib/spark/connector/lib/`.

S3 DistCp (s3-dist-cp)

Apache DistCp est un outil open source que vous pouvez utiliser pour copier de grandes quantités de données. S3 DistCp est similaire DistCp, mais optimisé pour fonctionner avec AWS, en particulier, Amazon S3. La commande pour S3 DistCp dans Amazon EMR version 4.0 et versions ultérieures est`s3-dist-cp`, que vous ajoutez en tant qu'étape dans un cluster ou sur la ligne de commande. Avec S3DistCp, vous pouvez copier efficacement de grandes quantités de données d'Amazon S3 vers HDFS où elles peuvent être traitées par les étapes suivantes dans votre cluster Amazon EMR. Vous pouvez également utiliser S3 DistCp pour copier des données entre des compartiments Amazon S3 ou de HDFS vers Amazon S3. S3 DistCp est plus évolutif et plus efficace pour copier en parallèle un grand nombre d'objets entre des compartiments et des AWS comptes.

Pour des commandes spécifiques démontrant la flexibilité de S3DistCP dans des scénarios réels, consultez Sept conseils d'utilisation de S3 DistCp sur le AWS blog Big Data.

Par exemple DistCp, S3 DistCp utilise MapReduce pour copier de manière distribuée. Il partage les tâches de copie, de gestion des erreurs, de récupération et de création de rapports entre plusieurs serveurs. Pour plus d'informations sur le projet DistCp open source Apache, consultez le DistCpguide dans la documentation d'Apache Hadoop.

Si S3 DistCp ne parvient pas à copier tout ou partie des fichiers spécifiés, l'étape du cluster échoue et renvoie un code d'erreur différent de zéro. Dans ce cas, S3 DistCp ne nettoie pas les fichiers partiellement copiés.

> **Important**
>
> S3 DistCp ne prend pas en charge les noms de compartiment Amazon S3 contenant le caractère de soulignement.
> S3 DistCp ne prend pas en charge la concaténation pour les fichiers Parquet. Utilisez PySpark plutôt. Pour plus d'informations, consultez Concaténation de fichiers parquet dans Amazon EMR.
> Pour éviter les erreurs de copie lorsque vous utilisez S3DistCp pour copier un seul fichier (au lieu d'un répertoire) de S3 vers HDFS, utilisez Amazon EMR version 5.33.0 ou ultérieure, ou Amazon EMR version 6.3.0 ou ultérieure.

DistCp Options S3

Bien que similaire à DistCp, S3 DistCp prend en charge un ensemble d'options différent pour modifier la façon dont il copie et compresse les données.

Lorsque vous appelez S3DistCp, vous pouvez spécifier les options décrites dans le tableau suivant. Les options sont ajoutées à l'étape à l'aide de la liste d'arguments. Des exemples d'DistCp arguments S3 sont présentés dans le tableau suivant.

Option	Description	Obligatoire
`--src=LOCATION`	Emplacement des données à copier. Cela peut être un emplacement dans HDFS ou Amazon S3. Exemple : `--src=s3://amzn-s3-demo-bucket/logs/j-3GYXXXXXX9IOJ/node` **Important** S3 DistCp ne prend pas en charge les noms de compartiment Amazon S3 contenant le caractère de soulignement.	Oui
`--dest=LOCATION`	Destination des données. Cela peut être un emplacement dans HDFS ou Amazon S3. Exemple : `--dest=hdfs:///output` **Important** S3 DistCp ne prend pas en charge les noms de compartiment Amazon S3 contenant le caractère de soulignement.	Oui
`--srcPattern=PATTERN`	Expression régulière qui filtre l'opération de copie sur un sous-ensemble des données à l'adresse `--src`. Si `--srcPattern` et `--groupBy` ne	Non

Option	Description	Obligatoire
	sont pas spécifiés, toutes les données à l'adresse `--src` sont copiées dans `--dest`. Si l'argument de l'expression régulière contient des caractères spéciaux, tels qu'un astérisque (*), l'expression régulière ou la chaîne `--args` complète doit être mise entre guillemets simples ('). Exemple : `--srcPattern=.*daemons.*-hadoop-.*`	

Option	Description	Obligatoire
`--groupBy=PATTERN`	Expression régulière qui oblige S3 DistCp à concaténer les fichiers correspondant à l'expression. Par exemple, vous pouvez utiliser cette option pour combiner tous les fichiers journaux écrits en une heure dans un fichier unique. Le nom de fichier concaténé est la valeur correspondante à l'expression régulière pour le regroupement. Les parenthèses indiquent comment les fichiers doivent être regroupés, en combinant tous les éléments qui correspondent à l'instruction entre parenthèses en un fichier de sortie unique. Si l'expression régulière n'inclut pas d'instruction entre parenthèses, le cluster échoue à l'DistCp étape S3 et renvoie une erreur. Si l'argument de l'expression régulière contient des caractères spéciaux, tels qu'un astérisque (*), l'expression régulière ou la chaîne `--args` complète doit être mise entre guillemets simples ('). Lorsque `--groupBy` est spécifié, seuls les fichiers qui correspondent au modèle spécifié sont copiés. Vous n'avez pas besoin de spécifier `--groupBy` et `--srcPattern` en même temps. Exemple : `--groupBy=.*subnetid.*([0-9]+-[0-9]+-[0-9]+-[0-9]+).*`	Non

Option	Description	Obligatoire
`--targetSize=SIZE`	Taille, en mébioctets (Mio), des fichiers à créer en fonction de l'option `--groupBy` . Cette valeur doit être un nombre entier. Lorsque cette option `--targetSize` est définie, S3 DistCp essaie de correspondre à cette taille ; la taille réelle des fichiers copiés peut être supérieure ou inférieure à cette valeur. Les travaux sont regroupés en fonction de la taille du fichier de données. Par conséquent, il est possible que la taille du fichier cible corresponde à la taille du fichier de données source. Si les fichiers concaténés par `--groupBy` sont plus grands que la valeur de `--targetSize` , ils sont divisés en fichiers plus petits, nommés de manière séquentielle avec une valeur numérique ajoutée à la fin de leur nom. Par exemple, un fichier concaténé `myfile.gz` peut être divisé en parties telles que : `myfile0.gz` , `myfile1.gz` , etc. Exemple : `--targetSize=2`	Non
`--appendToLastFile`	Spécifie le comportement de S3 DistCp lors de la copie vers des fichiers déjà présents depuis Amazon S3 vers HDFS. Il ajoute les données des nouveaux fichiers dans les fichiers existants . Si vous utilisez `--appendToLastFile` avec `--groupBy` , les nouvelles données sont ajoutées aux fichiers qui correspondent aux mêmes groupes. Cette option respecte également le comportement `--targetSize` lorsqu'il est utilisé avec `--groupBy.`	Non

Option	Description	Obligatoire
`--outputCodec=CODEC`	Spécifie le codec de compression à utiliser pour les fichiers copiés. Cela peut prendre les valeurs : `gzip`, `gz`, `lzo`, `snappy` ou `none`. Vous pouvez utiliser cette option, par exemple, pour convertir des fichiers d'entrée compressés avec Gzip en fichiers de sortie avec la compression LZO, ou pour décompresser les fichiers dans le cadre de l'opération de copie. Si vous choisissez un codec de sortie, le nom de fichier sera ajouté avec l'extension appropriée (par exemple, pour `gz` et `gzip`, l'extension est `.gz`). Si vous ne spécifiez pas de valeur pour `--outputCodec` , les fichiers sont copiés sans modification dans leur compression. Exemple : `--outputCodec=lzo`	Non
`--s3ServerSideEncryption`	Garantit que les données cibles sont transférées via SSL et automatiquement chiffrées dans Amazon S3 à l'aide d'une AWS clé côté service. Lorsque vous récupérez des données à l'aide de S3DistCp, les objets sont automatiquement déchiffrés. Si vous tentez de copier un objet non chiffré vers un compartiment Amazon S3 requérant le chiffrement, l'opération échoue. Pour plus d'informations, consultez Utilisation du chiffrement des données. Exemple : `--s3ServerSideEncryption`	Non

Option	Description	Obligatoire
`--deleteOnSuccess`	Si l'opération de copie est réussie, cette option oblige S3 DistCp à supprimer les fichiers copiés de l'emplacement source. Cela est utile si vous copiez des fichiers de sortie, tels que les fichiers journaux, d'un emplacement à un autre comme une tâche planifiée, et que vous ne voulez pas copier deux fois les mêmes fichiers. Exemple : `--deleteOnSuccess`	Non
`--disableMultipartUpload`	Désactive l'utilisation du téléchargement partitionné. Exemple : `--disableMultipartUpload`	Non
`--multipartUploadChunkSize=SIZE`	Taille, en Mio, de chaque partie d'un chargement partitionné sur Amazon S3. S3 DistCp utilise le téléchargement partitionné lorsqu'il copie des données dont la `multipartUploadChunkSize` taille est supérieure à. Pour améliorer la performance au travail, vous pouvez augmenter la taille de chaque pièce. La taille par défaut est de 128 Mio. Exemple : `--multipartUploadChunkSize=1000`	Non
`--numberFiles`	Ajoute aux fichiers de sortie des numéros séquentiels. Le compteur démarre à 0, à moins qu'une autre valeur soit spécifiée par `--startingIndex` . Exemple : `--numberFiles`	Non
`--startingIndex=INDEX`	Utilisé avec `--numberFiles` pour spécifier le premier numéro de la séquence. Exemple : `--startingIndex=1`	Non

Option	Description	Obligatoire
`--outputManifest=FILENAME`	Crée un fichier texte, compressé avec Gzip, qui contient une liste de tous les fichiers copiés par S3DistCp. Exemple : `--outputManifest=manifest-1.gz`	Non
`--previousManifest=PATH`	Lit un fichier manifeste créé lors d'un appel précédent à S3 à DistCp l'aide de l'`--outputManifest` indicateur. Lorsque l'`--previousManifest` indicateur est défini, S3 DistCp exclut les fichiers répertoriés dans le manifeste de l'opération de copie. Si `--outputManifest` est spécifié avec `--previousManifest`, les fichiers répertoriés dans le manifeste précédent apparaissent également dans le nouveau fichier manifeste, même si les fichiers ne sont pas copiés. Exemple : `--previousManifest=/usr/bin/manifest-1.gz`	Non
`--requirePreviousManifest`	Nécessite un manifeste créé lors d'un appel précédent à S3DistCp. Si la valeur spécifiée est false, aucune erreur n'est générée lorsqu'un manifeste précédent n'est pas spécifié. Par défaut, la valeur est true.	Non
`--copyFromManifest`	Inverse le comportement de S3 `--previousManifest` DistCp pour qu'il utilise le fichier manifeste spécifié comme liste de fichiers à copier, au lieu d'une liste de fichiers à exclure de la copie. Exemple : `--copyFromManifest --previousManifest=/usr/bin/manifest-1.gz`	Non

Option	Description	Obligatoire
`--s3Endpoint=ENDPOINT`	Spécifie le point de terminaison Amazon S3 à utiliser lors du chargement d'un fichier. Cette option définit le point de terminaison pour la source et la destination. S'il n'est pas défini, le point de terminaison par défaut est `s3.amazonaws.com` . Pour obtenir la liste des points de terminaison Amazon S3, consultez Régions et points de terminaison. Exemple : `--s3Endpoint=s3.eu-west-1.amazonaws.com`	Non
`--storageClass=CLASS`	La classe de stockage à utiliser lorsque la destination est Amazon S3. Les valeurs valides sont STANDARD et REDUCED_REDUNDANCY. Si cette option n'est pas spécifiée, S3 DistCp essaie de préserver la classe de stockage. Exemple : `--storageClass=STANDARD`	Non

Option	Description	Obligatoire
`--srcPrefixesFile=PATH`	fichier texte dans Amazon S3 (s3://), HDFS (hdfs:///) ou le système de fichiers local (file:/) qui contient une liste de préfixes `src`, un préfixe par ligne. S'il `srcPrefixesFile` est fourni, S3 ne DistCp listera pas le chemin src. Au lieu de cela, il génère une liste source en tant que résultat combiné de la liste de tous les préfixes spécifiés dans ce fichier. Le chemin d'accès relatif par rapport au chemin d'accès src, à la place de ces préfixes, sera utilisé pour générer les chemins de destination. Si `srcPattern` est également spécifié, il est appliqué aux résultats combinés de liste des préfixes sources pour filtrer encore plus les données d'entrée. Si `copyFromManifest` est utilisé, les objets dans le manifeste seront copiés et `srcPrefixesFile` sera ignoré. Exemple : `--srcPrefixesFile=PATH`	Non

Outre les options ci-dessus, S3 DistCp implémente l'interface Tool, ce qui signifie qu'elle prend en charge les options génériques.

Ajouter S3 DistCp en tant qu'étape dans un cluster

Vous pouvez appeler S3 DistCp en l'ajoutant en tant qu'étape dans votre cluster. Il est possible d'ajouter des étapes à un cluster lors de son lancement ou à un cluster en cours d'exécution à l'aide de la console, de l'interface de ligne de commande ou de l'API. Les exemples suivants illustrent l'ajout d'une DistCp étape S3 à un cluster en cours d'exécution. Pour plus d'informations sur l'ajout d'étapes dans un cluster, consultez Soumission de travail à un cluster dans le Guide de gestion d'Amazon EMR.

Pour ajouter une DistCp étape S3 à un cluster en cours d'exécution à l'aide du AWS CLI

Pour plus d'informations sur l'utilisation des commandes Amazon EMR dans le AWS CLI, consultez la référence des AWS CLI commandes.

- Pour ajouter une étape à un cluster qui appelle S3DistCp, transmettez les paramètres qui spécifient comment S3 DistCp doit effectuer l'opération de copie en tant qu'arguments.

  L'exemple suivant copie les journaux des démons de Amazon S3 dans `hdfs:///output`. Dans la commande suivante :

  - `--cluster-id` spécifie le cluster
  - `Jar`est l'emplacement du fichier DistCp JAR S3. Pour un exemple d'exécution d'une commande sur un cluster à l'aide de command-runner.jar, consultez Soumettre une étape JAR personnalisée pour exécuter un script ou une commande.
  - `Args`est une liste séparée par des virgules des paires nom-valeur des options à transmettre à S3. DistCp Pour obtenir la liste complète des options disponibles, consultez DistCp Options S3.

  Pour ajouter une étape de DistCp copie S3 à un cluster en cours d'exécution, placez ce qui suit dans un fichier JSON enregistré dans Amazon S3 ou dans votre système de fichiers local, comme dans cet *`myStep.json`* exemple. *`j-3GYXXXXXX9IOK`*Remplacez-le par votre ID de cluster et remplacez *`amzn-s3-demo-bucket`* par le nom de votre compartiment Amazon S3.

```
[
    {
        "Name":"S3DistCp step",
        "Args":["s3-dist-cp","--s3Endpoint=s3.amazonaws.com","--src=s3://amzn-s3-
demo-bucket/logs/j-3GYXXXXXX9IOJ/node/","--dest=hdfs:///output","--srcPattern=.*[a-
zA-Z,]+"],
        "ActionOnFailure":"CONTINUE",
        "Type":"CUSTOM_JAR",
        "Jar":"command-runner.jar"
    }
]
```

```
aws emr add-steps --cluster-id j-3GYXXXXXX9IOK --steps file://./myStep.json
```

Example Copie des fichiers journaux d'Amazon S3 vers HDFS

Cet exemple illustre également comment copier les fichiers journaux stockés dans un compartiment Amazon S3 vers HDFS en ajoutant une étape à un cluster en cours d'exécution. Dans cet exemple, l'option `--srcPattern` est utilisée pour limiter les données copiées dans les journaux des démons.

Pour copier les fichiers journaux d'Amazon S3 vers HDFS à l'aide de l'option `--srcPattern`, placez ce qui suit dans un fichier JSON enregistré dans Amazon S3 ou dans votre système de fichiers local en tant que *`myStep.json`* pour cet exemple. *`j-3GYXXXXXX9IOK`*Remplacez-le par votre ID de cluster et remplacez *`amzn-s3-demo-bucket`* par le nom de votre compartiment Amazon S3.

```
[
    {
        "Name":"S3DistCp step",
        "Args":["s3-dist-cp","--s3Endpoint=s3.amazonaws.com","--src=s3://amzn-s3-demo-
bucket/logs/j-3GYXXXXXX9IOJ/node/","--dest=hdfs:///output","--srcPattern=.*daemons.*-
hadoop-.*"],
        "ActionOnFailure":"CONTINUE",
        "Type":"CUSTOM_JAR",
        "Jar":"command-runner.jar"
    }
]
```

Nettoyage après l'échec de DistCp tâches S3

Si S3 DistCp ne parvient pas à copier certains ou tous les fichiers spécifiés, l'étape de commande ou de cluster échoue et renvoie un code d'erreur différent de zéro. Dans ce cas, S3 DistCp ne nettoie pas les fichiers partiellement copiés. Vous devez les supprimer manuellement.

Les fichiers partiellement copiés sont enregistrés dans le `tmp` répertoire HDFS dans des sous-répertoires avec l'identifiant unique de la tâche S3DistCp . Cet identifiant figure dans la sortie standard de la tâche.

Par exemple, pour une DistCp tâche S3 avec l'ID`4b1c37bb-91af-4391-aaf8-46a6067085a6`, vous pouvez vous connecter au nœud principal du cluster et exécuter la commande suivante pour afficher les fichiers de sortie associés à la tâche.

```
hdfs dfs -ls /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6/output
```

La commande renvoie une liste de fichiers similaires à la liste suivante :

```
Found 8 items
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/_SUCCESS
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00000
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00001
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00002
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00003
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00004
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00005
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00006
```

Vous pouvez ensuite exécuter la commande suivante pour supprimer le répertoire et l'ensemble du contenu.

```
hdfs dfs rm -rf /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6
```

Exécuter des commandes et des scripts sur un cluster Amazon EMR

Cette rubrique explique comment exécuter une commande ou un script en tant qu'étape sur votre cluster. L'exécution d'une commande ou d'un script sous forme d'étape est l'une des nombreuses méthodes permettant de soumettre du travail à un cluster. Elle est utile dans les situations suivantes :

- Lorsque vous n'avez pas accès SSH à votre cluster Amazon EMR
- Lorsque vous souhaitez exécuter une commande bash ou shell pour dépanner votre cluster

Vous pouvez exécuter un script soit lorsque vous créez un cluster, soit lorsque votre cluster est dans l'état `WAITING`. Pour exécuter un script avant le début du traitement de l'étape, utilisez plutôt une action d'amorçage. Pour plus d'informations sur les actions de démarrage, consultez la section Création d'actions d'amorçage pour l'installation de logiciels supplémentaires dans le Guide de gestion Amazon EMR.

Amazon EMR fournit les outils suivants pour vous aider à exécuter des scripts, des commandes et d'autres programmes intégrés au cluster. Vous pouvez appeler les deux outils à l'aide de la console de gestion Amazon EMR ou de l' AWS CLI.

`command-runner.jar`

Situé sur l'AMI Amazon EMR de votre cluster. Vous pouvez utiliser `command-runner.jar` pour exécuter des commandes sur votre cluster. Vous spécifiez `command-runner.jar` sans utiliser son chemin complet.

`script-runner.jar`

Hébergé sur Amazon S3 à `s3://<region>.elasticmapreduce/libs/script-runner/script-runner.jar` où *`<region>`* est la région dans laquelle réside votre cluster Amazon EMR. Vous pouvez utiliser `script-runner.jar` pour exécuter des scripts enregistrés localement ou sur Amazon S3 sur votre cluster. Vous devez spécifier l'URI complet de `script-runner.jar` lorsque vous soumettez une étape.

Soumettre une étape JAR personnalisée pour exécuter un script ou une commande

Les AWS CLI exemples suivants illustrent certains cas d'utilisation courants d'Amazon `command-runner.jar` EMR et `script-runner.jar` sur Amazon EMR.

Example : Exécution d'une commande sur un cluster à l'aide de **`command-runner.jar`**

Lorsque vous utilisez `command-runner.jar`, vous spécifiez des commandes, des options et des valeurs dans la liste d'arguments de votre étape.

L' AWS CLI exemple suivant envoie une étape à un cluster en cours d'exécution qui `command-runner.jar` invoque. La commande spécifiée dans la `Args` liste télécharge un script appelé *`my-script.sh`* depuis Amazon S3 dans le répertoire d'accueil de l'utilisateur hadoop. La commande modifie ensuite les autorisations du script et s'exécute*`my-script.sh`*.

Lorsque vous utilisez le AWS CLI, les éléments de votre `Args` liste doivent être séparés par des virgules, sans espace entre les éléments de la liste. Par exemple, `Args=[example-command,example-option,"example option value"]` plutôt que `Args=[example-command, example-option, "example option value"]`.

```
aws emr add-steps \
--cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Download a script from S3, change its permissions, and
 run it",ActionOnFailure=CONTINUE,Jar=command-runner.jar,Args=[bash,-c,"aws s3 cp s3://
amzn-s3-demo-bucket/my-script.sh /home/hadoop; chmod u+x /home/hadoop/my-script.sh;
 cd /home/hadoop; ./my-script.sh"]
```

Example : Exécution d'un script sur un cluster à l'aide de **`script-runner.jar`**

Lorsque vous utilisez `script-runner.jar`, vous spécifiez le script que vous souhaitez exécuter dans la liste des arguments de votre étape.

L' AWS CLI exemple suivant envoie une étape à un cluster en cours d'exécution qui `script-runner.jar` invoque. Dans ce cas, le script appelé *`my-script.sh`* est stocké sur Amazon S3. Vous pouvez également spécifier des scripts locaux qui sont stockés sur le nœud principal de votre cluster.

```
aws emr add-steps \
--cluster-id j-2AXXXXXXGAPLF \
```

```
--steps Type=CUSTOM_JAR,Name="Run a script from S3 with script-
runner.jar",ActionOnFailure=CONTINUE,Jar=s3://us-west-2.elasticmapreduce/libs/script-
runner/script-runner.jar,Args=[s3://amzn-s3-demo-bucket/my-script.sh]
```

Autres moyens d'utiliser `command-runner.jar`

Vous pouvez également utiliser `command-runner.jar` pour soumettre du travail à un cluster à l'aide d'outils tels que `spark-submit` ou `hadoop-streaming`. Lorsque vous lancez une application à l'aide de `command-runner.jar`, vous spécifiez `CUSTOM_JAR` comme type d'étape au lieu d'utiliser une valeur telle que `SPARK`, `STREAMING` ou `PIG`. La disponibilité des outils varie en fonction des applications que vous avez installées sur le cluster.

L'exemple de commande suivant utilise `command-runner.jar` pour soumettre une étape à l'aide de `spark-submit`. La `Args` liste indique `spark-submit` la commande, suivie de l'URI Amazon S3 de l'application Spark *`my-app.py`* avec les arguments et les valeurs.

```
aws emr add-steps \
--cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Run spark-submit using command-
runner.jar",ActionOnFailure=CONTINUE,Jar=command-runner.jar,Args=[spark-submit,S3://
amzn-s3-demo-bucket/my-app.py,ArgName1,ArgValue1,ArgName2,ArgValue2]
```

Le tableau suivant identifie les outils supplémentaires que vous pouvez exécuter à l'aide de `command-runner.jar`.

Nom de l'outil	Description
`hadoop-streaming`	Soumet un programme de streaming Hadoop. Dans la console et dans d'autres SDKs, il s'agit d'une étape de streaming.
`hive-script`	Exécute un script Hive. Dans la console et SDKs, il s'agit d'une étape Hive.
`pig-script`	Exécute un script Pig. Dans la console SDKs, il s'agit d'un Pig step.
`spark-submit`	Exécute une application Spark. Dans la console, il s'agit d'une étape Spark.

Nom de l'outil	Description
`hadoop-lzo`	Exécute l'indexeur Hadoop LZO sur un répertoire.
`s3-dist-cp`	Effectuez la copie distribuée de grandes quantités de données d'Amazon S3 dans HDFS. Pour de plus amples informations, veuillez consulter S3 DistCp (s3-dist-cp).

AWS Glossaire

Pour la AWS terminologie la plus récente, consultez le AWS glossaire dans la Glossaire AWS référence.

Les traductions sont fournies par des outils de traduction automatique. En cas de conflit entre le contenu d'une traduction et celui de la version originale en anglais, la version anglaise prévaudra.